

ATTN: Ed Hart. 66252

Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Celine Qian Examiner #: 75710 Date: 5/8/02
 Art Unit: 1636 Phone Number 30 6-0283 Serial Number: 09/662128
 Mail Box and Bldg/Room Location: 11E.05 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Modified cre recombinase gene for mammals
 Inventors (please provide full names): Shuji Hiyagawa, Masaru Okabe

Earliest Priority Filing Date: 9/17/1999

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

search SEQ ID NO: 1 (both commercial & interference)

search nucleic acid sequence encoding amino acid
 sequence of SEQ ID NO: 2

RECEIVED
 MAY - 8 2002
 (STIC)

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) <u>1</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>5/8/02</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>5/14/02</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>05</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2002, 00:54:43 ; Search time 253.13 Seconds
(without alignments)
7121.875 Million cell updates/sec

Title: US-09-662-128A-1
Perfect score: 1050
Sequence: 1 atgcccaagaagaagagaa.....gcctgctggaagcagcgac 1050

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1050	100.0	22	AAH41175	Mammalian Cre reco
2	755	71.9	22	AAH41175	Chimeric Cre reco
3	754.4	71.8	22	AAH41175	Chimeric Cre reco
4	639.4	60.9	22	AAH41175	Chimeric Cre reco
5	639.4	60.9	22	AAH41175	Chimeric Cre reco
6	639.4	60.9	22	AAH41175	Chimeric Cre reco
7	636.4	60.6	22	AAH41175	Chimeric Cre reco
8	630	60.0	22	AAH41175	Chimeric Cre reco
9	621.8	59.2	21	AAH41175	Chimeric Cre reco

10	617.4	58.8	1074	22	AAH41175	TAT-Cre fusion DNA
11	617.4	58.8	1623	22	AAH41175	Delta VP22Cre-Stre
12	617.4	58.8	1740	21	AAH41175	GST-Cre coding reg
13	617.4	58.8	2004	18	AAH41175	Chimeric Cre-LDB-G
14	617.4	58.8	2004	18	AAH41175	VP22-Cre fusion DN
15	617.4	58.8	2055	22	AAH41175	VP22CreStreptag fu
16	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
17	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
18	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
19	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
20	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
21	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
22	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
23	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
24	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
25	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
26	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
27	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
28	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
29	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
30	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
31	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
32	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
33	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
34	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
35	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
36	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
37	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
38	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
39	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
40	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
41	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
42	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
43	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
44	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G
45	617.4	58.8	2346	22	AAH41175	Chimeric Cre-LDB-G

ALIGNMENTS

RESULT 1	
AAH41175	
ID	AAH41175 standard; DNA; 1050 BP.
AC	AAH41175;
DT	23-AUG-2001 (first entry)
DE	Mammalian Cre recombinase gene.
KW	Mammalian: Cre recombinase; organ transplantation; gene therapy;
KW	animal disease model; ds.
OS	Mammalia.
PN	JP2001086989-A.
PD	03-APR-2001.
PF	17-SEP-1999; 99JP-0264364.
PR	17-SEP-1999; 99JP-0264364.
PA	(OSAU) UNIV OSAKA.
DR	WPI; 2001-360321/38.
DR	P-PSDB; AAB98695.
PT	Mammalian type Cre recombinase gene modified so as to have a high
PT	expression efficiency in a mammal by selecting a codon of high
PT	frequency for use in the mammal for expressing Cre recombinase -


```

Db      1457  cgcgcgcagatatacgaatcgtgcatctctggtggtgattgctatatacaccctgttacct 1516
QY      541  atcgccagatcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Db      1517  atagccgaatctgcagacgacgaggttaagatatctcagctacgtgagtgaggagatg 1576
QY      601  ctgacacacatcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 660
Db      1577  ttaatccatattgacagacgaacgcgtgttagcacgcgcgcgcgcgcgcgcgcgcgcgcgc 1636
QY      661  agccttgagcgtgacacgctggtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
Db      1637  agcctggggggaactaaacggtctgcagcagatcttcgctccctggtgtagctgtagt 1696
QY      721  cccaacaactactgtcttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780
Db      1697  cgaataactactcgttcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1756
QY      781  agccagctgagacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 840
Db      1757  agccagctatcaactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1816
QY      841  agcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 900
Db      1817  ggcgcgtaagatgactcgtgctcagagatacctgctgctcgtgacacagctgcccgtgtc 1876
QY      901  ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
Db      1877  ggaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1936
QY      961  ggcctggacacgaatgtaacatcgtaagacacacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1020
Db      1937  ggcctggacacgaatgtaacatcgtaagacacacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1996
QY      1021  gccatggtgcccgtcgtgagagacgcga 1049
Db      1997  gcaatggtgcccgtcgtgagagatgcga 2025

RESULT  6
AAD04928  standard; DNA; 5365 BP.
XX
AC      AAD04928;
XX
DT      17-JUL-2001 (first entry)
XX
DE      Retroviral vector PBABE-pgkCre used as gene trap in embryonic stem cells.
XX
KW      Gene trapping construct; conditional mutation; unidirectional inversion;
KW      recombinase recognition sequence; RRS; disruption cassette;
KW      selection cassette; transgenic organism; retroviral vector;
KW      PBABE-pgkCre; cre recombinase; phosphoglycerate kinase; PGK promoter; ds.
XX
OS      Chimeric - Moloney murine leukemia virus.
OS      Chimeric - Mus sp.
OS      Chimeric - Bacteriophage P1.
OS      Chimeric - Rhesus macaque polyoma virus.
XX
FH      Key Location/Qualifiers
FT      LTR 8..480
FT      /*tag= a
FT      /note= "5' partial MMLV U3, MMLV R and MMLV U5"
FT      /partial
FT      8..335
FT      /*tag= b
FT      /note= "5' partial MMLV U3"
FT      /partial
FT      336..402
FT      /*tag= c
FT      /rpt_type= DIRECT
FT      /note= "5' MMLV R direct repeat"
XX

```

```

FT      misc_feature 403..480
FT      /*tag= d
FT      /note= "5' MMLV U5"
FT      481..1374
FT      /*tag= e
FT      /note= "MMLV primer binding site and extended
FT      packaging signal"
FT      1417..1921
FT      /*tag= f
FT      /note= "Mouse phosphoglycerate kinase (PGK) promoter"
FT      1972..3024
FT      CDS
FT      /*tag= g
FT      /product= "Bacteriophage P1 Cre recombinase protein with
FT      a simian virus 40 large T-antigen N-terminal nuclear
FT      localisation signal"
FT      3088..3168
FT      /*tag= h
FT      /note= "Promoter/enhancer deleted MMLV U3"
FT      3187..3332
FT      LTR
FT      /*tag= i
FT      /note= "3' MMLV R and MMLV U5"
FT      3187..3253
FT      repeat_unit
FT      /*tag= j
FT      /rpt_type= DIRECT
FT      /note= "3' MMLV R direct repeat"
FT      3254..3332
FT      misc_feature
FT      /*tag= k
FT      /note= "3' MMLV U5"
XX
PN      WO200129208-A1.
XX
PD      26-APR-2001.
XX
XX      16-OCT-2000; 2000WO-EP10162.
XX
PR      16-OCT-1999; 99EP-0120592.
PR      27-OCT-1999; 99US-0162016.
XX
PA      (ARTE-) ARTEMIS PHARM GMBH.
PA      (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX
PI      Kuehn R, Von Melchener H, Altschmied J;
XX      WPI; 2001-308486/32.
XX
DR      New gene trapping construct capable of causing conditional mutations in
PT      genes, comprises functional DNA segment inserted in sense or antisense
PT      direction relative to gene to be trapped
XX
PS      Example 1; Page 55-57; 78pp; English.
XX
XX      The present invention relates to a conditional gene trapping construct
XX      capable of causing conditional mutations in genes. The gene trapping
XX      construct comprises two functional DNA segments, each being flanked by
XX      two recombinase recognition sequences (RRSs) specific to site specific
XX      recombinase which is capable of unidirectional inversion of double
XX      standard DNA segment. One of the DNA segment (disruption cassette) is
XX      inserted in antisense orientation relative to the transcriptional
XX      orientation of the gene to be trapped. The other DNA segment (selection
XX      cassette) is inserted in sense direction relative to the transcriptional
XX      orientation of the gene to be trapped. The cell comprising the gene
XX      trapping construct is useful for the identification and/or isolation of
XX      genes. The transgenic organism comprising the gene trapping construct is
XX      useful to study gene function at various developmental stages. The gene
XX      trapping construct is useful for mutationally inactivating all cellular
XX      genes. The present sequence is retroviral vector PBABE-pgkCre which is
XX      used as a gene trap in embryonic stem (ES) cells. The vector
XX      PBABE-pgkCre is constructed by inserting Bacteriophage P1 cre recombinase
XX      coding region with a simian virus 40 large T-antigen N-terminal nuclear
XX      localisation signal and mouse phosphoglycerate kinase (PGK) promoter into
XX      Moloney murine leukemia virus based vector PBABEpuro.
XX

```

Sequence 5365 BP; 1208 A; 1448 C; 1402 G; 1307 T; 0 other;

Query Match 60.9%; Score 639.4; DB 22; Length 5365;
Best Local Similarity 75.6%; Pred. No. 1.8e-85;
Matches 793; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

```

QY 1 atgcccaagaagaagatgagcaaccctgacccgacacgaacacccgcgc 60
DB 1972 atgcccaagaagaagatgagcaaccctgacccgacacgaacacccgcgc 60
QY 61 ctgcccgtgagcagcagcagcagcagcagcagcagcagcagcagcagc 120
DB 2032 ttaccggtcagatgagcagcagcagcagcagcagcagcagcagcagc 120
QY 121 cgcagagccttcagcagcagcagcagcagcagcagcagcagcagcagc 180
DB 2092 cgcagagccttcagcagcagcagcagcagcagcagcagcagcagcagc 180
QY 181 gctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
DB 2152 gcatgtgcaagtgatgaatacgcgaagatgttccgcgagacatgagcagc 240
QY 241 taacctgctacccgtcagcagcagcagcagcagcagcagcagcagcagc 300
DB 2212 taacctgctacccgtcagcagcagcagcagcagcagcagcagcagcagc 300
QY 301 cagctgacacatgctgacagcagcagcagcagcagcagcagcagcagcagc 360
DB 2272 cagctgacacatgctgacagcagcagcagcagcagcagcagcagcagcagc 360
QY 361 agctctgtatgctgacagcagcagcagcagcagcagcagcagcagcagc 420
DB 2332 tcaactgtgtatgctgacagcagcagcagcagcagcagcagcagcagcagc 420
QY 421 gacctggtcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
DB 2392 gacctggtcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
QY 481 cgtctgacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540
DB 2452 cgtctgacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540
QY 541 atccgcagatcgccgcagcagcagcagcagcagcagcagcagcagcagcagc 600
DB 2512 atccgcagatcgccgcagcagcagcagcagcagcagcagcagcagcagcagc 600
QY 601 ctgattcacatcgccgcagcagcagcagcagcagcagcagcagcagcagcagc 660
DB 2572 ctgattcacatcgccgcagcagcagcagcagcagcagcagcagcagcagcagc 660
QY 661 agcctggtcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
DB 2632 agcctggtcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
QY 721 cccaacacacacacacacacacacacacacacacacacacacacacacacac 780
DB 2692 cccaacacacacacacacacacacacacacacacacacacacacacacacac 780
QY 781 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
DB 2752 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
QY 841 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900
DB 2812 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900
QY 901 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
DB 2872 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
QY 961 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020
DB 961 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020

```

```

DB 2932 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2991
QY 1021 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1049
DB 2992 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3020

```

RESULT 7
AAK19900
ID AAK19900 standard; DNA; 2293 BP.

AC AAK19900;
DT 11-JUN-1999 (first entry)

DE Plasmid pOC304M containing mPl promoter and Cre coding sequences.

KW Mouse; mPl: promoter; site-specific recombination; Cre recombinase;
KM protamine 1; fusion gene; transgenic mouse; ss.

OS Synthetic.
OS Mus musculus.

PN MO9910488-A1.

PD 04-MAR-1999.

PF 28-AUG-1998; 98MO-US17852.

PR 28-AUG-1997; 97US-0919501.

PA (SALK) SALK INST BIOLOGICAL STUDIES.

PI O'Gorman S, Wahl G;

DR WPI: 1999-190617/16.

PT New nucleic acid construct, comprising a germline-specific promoter
operably linked to a recombinase coding sequence - useful for
producing subtle, conditional or tissue-specific mutations

PS Example 1; Page 48; 54pp; English.

The present invention describes a nucleic acid construct, comprising a
germline-specific promoter operably linked to a recombinase coding
sequence. The construct can be useful for modulating agronomic
characteristics, including (a) biotic stress tolerance, increased
resistance to herbicides, pest damage, and viral, bacterial, and fungal
diseases, improvement of crop nutritional quality, reduction of
post-harvest losses, improved quantity and composition of endogenous
compounds, and production of new plant-derived compounds. The
production of alleles containing single recombinase target sites and a
desired mutation are useful for producing subtle and conditional
mutations that require generation of alleles with minimal structural
alterations. They are also useful for generating null alleles.
mPl-mediated nucleic acid constructs are useful for producing embryos
expressed at high levels in the germ line, but not to a functionally
significant degree in ES cells or embryonic or adult somatic tissue.
This enables modulation of recombinase target nucleic acid sequences in
the early embryo. In addition, tissue-specific and conditional
recombinase tools permit the generation of transgenic plants and
animals. The present sequence represents a plasmid containing mouse
protamine 1 promoter and Cre coding sequences from an example of the
present invention.

Sequence 2293 BP; 555 A; 559 C; 603 G; 576 T; 0 other;

Query Match 60.6%; Score 636.4; DB 20; Length 2293;
Best Local Similarity 75.5%; Pred. No. 5.3e-85;
Matches 790; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

[illegible]

Best Local Similarity 75.1%; Pred. No. 4, 1e-84;
Matches 786; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

```

OY 4 cccaagaagaaggatgagcaacctgtcagccgtgacacgaacctgcccgcctg 63
Db 1201 ccaagaagaagaaggatcttcgaattactgaccgtacacaaatttgcctgcat 1260
OY 64 cccgtgagacgacacgagcagatgagcgaacccgattgagatgttccgcgacgc 123
Db 1261 ccggtcgtatgcaacgagtgatgaggttcgcaagaacccgagatgatagttaaggatcgc 1320
OY 124 cagggcctcagcagacacacctggaagatgtcgtcagcgtgttcgcgcgcgcgc 183
Db 1321 cagggcgtttctcgaacatacctggaanaatgtcttcgtcgttgcggttcgtggcgagca 1380
OY 184 tggtagcaagctgaacacacgagcagatggttcccgccgagcccgagagcgtgcgactac 243
Db 1381 tggtagcaagctgaataacacggaatggttcccgccgagacactggaagatgttcgcatat 1440
OY 244 ctgctgtacctgacagggccgcgcgttgcgcgtggaagacatccagcagaccttgagccag 303
Db 1441 ctctcatatcttcagcgcgcgcgttcgtgcagtaaaactatccagacacatttgggcag 1500
OY 304 ctgaacatctgacgacgacgagcgcgcgttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 363
Db 1501 ctgaacatgcttcacatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1560
OY 364 ctggtgtatgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 423
Db 1561 ctggtgtatgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1620
OY 424 ctggtgtatgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 483
Db 1621 ctggtgtatgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1680
OY 484 tgcgaagacatccgcaacacctggtcctcctgtgacatccgcaacacacccctgcgcgcgcgc 543
Db 1681 tgcgaagatatacgttaacttcgtcgttcgtggtatgtctataacacccctgtaacgtata 1740
OY 544 ggcgaagatccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 603
Db 1741 ggcgaagatccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1800
OY 604 atccacatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 663
Db 1801 atccacatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1860
OY 664 ctggggtgacacgaagctggtgtgagcgcgtgacacgcgtgagcgcgttgcgcgcgcgcgcgcgc 723
Db 1861 ctggggtgacacgaagctggtgtgagcgcgtgacacgcgttgcgcgcgcgcgcgcgcgcgcgcgcgc 1920
OY 724 aacaaactactgtcttcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 783
Db 1921 aacaaactactgtcttcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1980
OY 784 caactcagcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 843
Db 1981 caactcagcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2040
OY 844 gccaagaagacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 903
Db 2041 gccaagaagacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2100
OY 904 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 963
Db 2101 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2160
OY 964 tggagcaacgltgacacatcgtgtgatacctacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1023
Db 2161 tggagcaacgltgacacatcgtgtgatacctacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2220
OY 1024 atggtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1049

```

Db 2221 atggtgc 2246

RESULT 9

AA62541
ID AAC62541 standard; DNA: 1172 BP.

XX AAC62541;

XX 07-FEB-2001 (first entry)

XX Cre wild-type coding sequence.

XX Cre variant recognition site; lox site; recombinase;

KW variant recombination site; hybrid crop production; seedless crop;

KW phage packaging; cloning; ds.

OS Unidentified.

XX WO20060091-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US09154.

XX 06-APR-1999; 99US-0127977.

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

PI Sauer BL, Rufer AW;

XX WPI: 2000-665010/64.

XX Identifying variant recombinases mediating recombination at variant

PT sites (VNS) by contacting a mutant recombinase, a first and second VNS

PT having a reporter gene, and a second nucleic acid having 2 VNS and a

PT reporter gene

XX Example 1; Page 98-101; 144pp; English.

XX The present invention relates to the identification of recombinase
CC variants which have an altered specificity. They are tested using
CC constructs containing variant recognition sites, which are not recognised
CC by non-mutant recombinase but undergo recombination in the presence of a
CC variant enzyme. Variant recombinases are useful in the production of a
CC genetically modified crop plants, particularly seedless varieties, and in
CC phage packaging, which has uses in cloning.

XX Sequence 1172 BP; 299 A; 263 C; 323 G; 287 T; 0 other;

Query Match 59.2%; Score 621.8; DB 21; Length 1172;

Best Local Similarity 74.5%; Pred. No. 7, 7e-83;

Matches 782; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

OY 1 atcccaagaagaagaaggatgagcaacctgtcagccgtgacacgaacctgcccgcgc 60

Db 35 atccctagacgtgagtgtaaatgtccaatctactgacgcgtacacaaatttgcctgcga 94

OY 61 ctgcccgttgaagccacacgc 120

Db 95 ttaccggtcgaatgcaacgagtgatgaggttcgcaagaacctgatatgacatgttcaaggat 154

OY 121 cgcgaagcctcagcagacacacctggaagatgctgagcgcgtgtgcgcgcgcgcgcgcgcgcgcgc 180

Db 155 cgc 214

OY 181 gctcgtgtgaagcctgaacacacgcgaagtgttcccgccgagccgcgcgcgcgcgcgcgcgcgcgcgc 240

Db 215 gatcgtgtgaagcttgaataacgcgaatgcttcccgccgagacactgtaagtgttcgcgat 274

OY 241 tacctgtctactcctgacggccgc 300

[illegible]

(ARTE-) ARTEMIS PHARM GMBH.
Schwenk F;
WPI: 2001-441873/47.
P-PSTDB; AAE05270.

Using site-specific DNA recombinase domain/protein transduction domain fusion proteins for inducing target gene alterations in organisms or cell cultures -

Claim 17; Page 44-46; 85pp; English.

The present invention relates to use of fusion proteins comprising a site-specific DNA recombinase domain e.g. Cre and a protein transduction domain (PTD) e.g. the Human immunodeficiency virus (HIV) derived TAT peptide, for preparing an agent for inducing target gene alterations in a living organism or cell culture. The present invention also provides a method for inducing gene alteration in living organisms using the fusion proteins of the invention. The present sequence is a DNA encoding delta VP22Cre-StrepTag fusion protein. The VP22 sequence is from Human spumaretrovirus (HSV).

Sequence 1623 BP; 368 A; 440 C; 471 G; 344 T; 0 other:

[illegible]

Db 1142 aacgcgttagacccgagcgttagagagagcacttagcctggtgggtaactaaatcgg 1201
 Qy 683 tggagcgttagatcagcgttagagcgttagcagcagcccaactactctctgc 742
 Db 1202 tcgagcgttagatcctcgtctcgtgtagcgtatgatccgaataactacccgtttgc 1261
 Qy 743 gcgtgcagaaagacggtggtgcgcgcgcagccagccagcagcagcgtgagccggcc 802
 Db 1262 gggtcagaaaatggtgttcgcgcgcagcactgcacacagccagcactaactcggccc 1321
 Qy 803 tggagcgttagccttgcagcagccacccagcgttagcagcagcagcagcagcagc 862
 Db 1322 tggagcgttagccttgcagcagccacccagcgttagcagcagcagcagcagcagc 1381
 Qy 863 agcgttagccttgcagcagccacccagcgttagcagcagcagcagcagcagcagc 922
 Db 1382 agagatcctgctgctgtctgtagacagcgtccggtgtcggagcgcgcagatagccc 1441
 Qy 923 ggcgcgcgttagcagcagccagcagcagcagcagcagcagcagcagcagcagcagc 982
 Db 1442 ggcgcgcgttagcagcagccagcagcagcagcagcagcagcagcagcagcagcagc 1501
 Qy 983 tgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1042
 Db 1502 tcatgaactatcctcgttagcagcagcagcagcagcagcagcagcagcagcagcagc 1561
 Qy 1043 acgcgcga 1049
 Db 1562 atgcgcga 1568
 RESULT 12
 AAZ58065
 ID AAZ58065 standard; DNA; 1740 BP.
 AC AAZ58065;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE GST-Cre coding region within pOL123.
 XX
 KW Cre recombinase; loxP; site-specific recombination; subcloning;
 KW Unifector Fusion System; Unifector Plasmid-Fusion System;
 KW phage P1; plasmid pOL123; ss.
 XX
 OS Chimeric - Unidentified.
 OS Chimeric - Bacteriophage hp1.
 OS
 PN WO200005355-A1.
 XX
 PD 03-FEB-2000.
 XX
 PF 27-JUL-1998: 98WO-US15481.
 XX
 PR 24-JUL-1998: 98US-0122384.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 PI Elledge SJ, Liu Q;
 DR WPI; 2000-171429/15.
 DR P-PSDB; AAY58676.
 XX
 PT New methods for recombination of nucleic acid constructs, useful for
 XX the rapid subcloning of nucleic acid sequences in vivo and in vitro
 XX
 PS Example 3; Page 67-68; 110pp; English.
 CC The present sequence is that of the coding region within plasmid
 CC pOL123 encoding a Gst-Cre fusion protein (see AAY58676). Cre is a
 CC site-specific recombinase of phage P1 that catalyzes recombination
 CC of DNA between 2 loxP sites. The cre gene was isolated by PCR from

CC plasmid pBS39, cloned into a TA cloning vector and then subcloned
 CC into pGEX-2TKcs to generate pOL123. The Gst-Cre fusion protein was
 CC expressed in Escherichia coli BL21 (DE3) transformants. The
 CC recombinant protein provides a source of purified Cre recombinase
 CC for use in the in vitro recombination of plasmids. The invention
 CC provides novel Unifectors or PUN1 vectors (see AA58064) that have a
 CC sequence-specific recombinase target site (e.g. loxP) preceding the
 CC insertion site for the gene of interest, a selectable marker gene
 CC (optional), and a conditional origin of replication that is active
 CC only in host cells expressing the requisite transacting replication
 CC factor (optional). The vectors are used in a novel method for the
 CC rapid subcloning of nucleic acid sequences in vivo or in vitro
 CC without the need of restriction enzymes. The method is referred to
 CC as the Unifector Fusion System or Unifector Plasmid-Fusion System
 CC (UPS). The UPS uses a site-specific recombinase to catalyze plasmid
 CC fusion between a univector and host vectors containing regulatory
 CC information. In some embodiments, plasmid fusion events are
 CC genetically selected and result in placement of the gene of interest
 CC under the control of novel regulatory elements. A second UPS-related
 CC method allows for the precise transfer of coding sequences alone from
 CC a Unifector into a host vector. UPS further provides means for the
 CC subcloning of entire nucleic acid libraries and the directional
 CC cloning of linear nucleic acid molecules, e.g. PCR products.

Sequence 1740 BP; 464 A; 362 C; 443 G; 471 T; 0 other;

Query Match 58.8%; Score 617.4; DB 21; Length 1740;
 Best Local Similarity 75.1%; Pred. No. 3,2e-82;
 Matches 771; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

Qy 23 tgaagcagcctctacacgttagcagcagcagcagcagcagcagcagcagcagcagcagc 82
 Db 710 tggcgaattactacacgttagcagcagcagcagcagcagcagcagcagcagcagcagc 769
 Qy 83 acgaagtgccgaagaacctgtagacatgttccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 142
 Db 770 atgagtgccgaagaacctgtagacatgttccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 829
 Qy 143 cctggaagatgtcgttagcagcagcagcagcagcagcagcagcagcagcagcagcagc 202
 Db 830 cctggaagatgtcgttagcagcagcagcagcagcagcagcagcagcagcagcagcagc 889
 Qy 203 gaaatgtgttccgc 262
 Db 890 ggaatgtgttccgc 949
 Qy 263 ggcgcgtgctgtagaagacatccagcagcagcagcagcagcagcagcagcagcagcagc 322
 Db 950 ggcgcgtgctgtagaagacatccagcagcagcagcagcagcagcagcagcagcagcagc 1009
 Qy 323 gcaagcgtgctgtagaagacatccagcagcagcagcagcagcagcagcagcagcagcagc 382
 Db 1010 gtlccggtgctgtagaagacatccagcagcagcagcagcagcagcagcagcagcagcagc 1069
 Qy 383 gcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 442
 Db 1070 gaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1129
 Qy 443 actgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 502
 Db 1130 attgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1189
 Qy 503 tggcgttctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 562
 Db 1190 tggcgttctgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1249
 Qy 563 ggcgtgaagacatccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 622
 Db 1250 ggcgtgaagacatccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1309
 Qy 623 agacactgtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 682

RESULT	13
AAV05702	
ID	AAV05702 standard; cDNA; 2004 BP.
XX	

FT	location/vallieres	1..2004	/*tag= a	/product= "Cre-LDB-hGR(1747T) fusion protein"	FT
FT		1..1029	/*tag= b		FT
FT		1030..1131	/label= A	/note= "derived from bacteriophage PI Cre gene"	FT
FT			/*tag= c		FT
FT			/label= B		FT
FT			/note= "derived from C-terminus of human glucocorticoid receptor D region"		FT
FT		1132..1869	/*tag= d		FT
FT			/label= C		FT
FT		1876..2001	/note= "derived from LDB-hGR(1747T) sequence"		FT
FT			/*tag= e		FT
FT			/label= D		FT
FT			/note= "derived from region F of human oestrogen receptor sequence"		FT
XX					XX
PN	FR2745008-A1.				PN
XX					XX

[illegible]

```

Db 482 tggcattctcgggagattgttataaacaccctgttcaatagtcgaaattgcagatca 541
OY 563 gcttgaagagacatcagccgacacgagcgccgcatgtatccacatcggcgacaca 622
Db 542 gggttaaagatatccacactgactgacggtgggagatgttaaccatattgacagaaga 601
OY 623 agaccctgtgtgacacccgacggtgaggaagagccctgtgacgttggcgctgacgaatcgg 682
Db 602 aaacgtgtgttagcacccgacaggtgttagaagaagcacttaagccttgggggtaactaaatcgg 661
OY 683 tggagcgtcgtgacatcagcgtgtgagcgctgtgacgacgaccccaacacatacctgtctgc 742
Db 662 tggagcgtgagattctccgtctgtgtgtgtagctgtagatccgataactaactgttttgc 721
OY 743 gctgtgcgaagaagcggtgtgacgccccacgacgacgacgacgacgtgtgacacggcgcc 802
Db 722 gggtcagaanaaatgtgtgtgacgacacatcgtcgaacgacgacgacgacatcaactcggccc 781
OY 803 tggagggcattcttcgagcgcacccacacgctgtatctagcgcgcgaaggaacgacgacgccc 862
Db 782 tggagggagatttttgaagaacatcatcgtattttagcgcgctaaagatgactgtctgtc 841
OY 863 agcgtacctgtgctgtgagcgacgacgacgacgacgacgacgacgacgacgacgacgacgac 922
Db 842 agagatacctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 901
OY 923 ggcgcgctgtgacatcccgagatcatcagcgcgctgtgacgacgacgacgacgacgacgacgac 982
Db 902 ggcgtgtgagtttcaatacaggaagatcatcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 961
OY 983 tgtgtacatacatcgcgaacacgtgagacgacgacgacgacgacgacgacgacgacgacgacgac 1042
Db 962 tcatgacatacatcgcgaacacgtgagacgacgacgacgacgacgacgacgacgacgacgacgac 1021
OY 1043 acggcgca 1049
Db 1022 atggcgca 1028

```

RESULT 14
 AAD09259
 ID AAD09259 standard; DNA; 2004 BP.
 AC AAD09259;
 DT 12-SEP-2001 (first entry)
 DE VP22-Cre fusion DNA.
 KW DNA recombinase domain; protein transduction domain; PTD;
 KW gene alteration; VP22-Cre fusion protein; Human immunodeficiency virus;
 KW HIV; Human spumaretrovirus; HSV; ds.
 OS Chimeric - Human spumaretrovirus.
 OS Chimeric - Unidentified.
 FT Key Location/Qualifiers
 FT CDS 1..2004
 FT /product= "VP22-Cre fusion protein"
 PN WO200149832-A2.
 PD 12-JUL-2001.
 PF 05-JAN-2001; 2001MO-EP00060.
 PR 07-JAN-2000; 2000EP-0100351.
 PR 10-NOV-2000; 2000EP-0124595.
 PA (ARTE-) ARTEMIS PHARM GMBH.

```

PI Schwenk F;
XX WPI; 2001-441873/47.
DR P-PSDB; AAE05266.
XX Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PT cell cultures -
XX Claim 17; Page 32-35; 85pp; English.
PS
XX The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived TAR peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a DNA encoding VP22-Cre fusion
CC protein. The VP22 sequence is from Human spumaretrovirus (HSV).
XX
XX Sequence 2004 BP; 419 A; 608 C; 592 G; 385 T; 0 other;
SQ

```

Query Match 58.8%; Score 617.4; DB 22; Length 2004;
 Best Local Similarity 75.1%; Pred. No. 3.2e-62;
 Matches 771; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

```

OY 23 tgaagcaactgtgtacacgtgacacgacacgtgacacgacgacgacgacgacgacgacgacgacgac 82
Db 974 tgtccaatttcaatcgtacgtacacccaataattgtcgtcattacacggtgtgtgtgtgtgtgtgt 1033
OY 83 aagaagtgcgaagaacacgtgtgacatgttcgagacgacgacgacgacgacgacgacgacgacgacgac 142
Db 1024 atgaggttcgaagaacacgtgtgacatgttcgagacgacgacgacgacgacgacgacgacgacgacgac 1093
OY 143 cctggaagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 202
Db 1094 cctggaagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1153
OY 203 gcaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 262
Db 1154 ggaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1213
OY 263 ggcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 322
Db 1214 ggcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1273
OY 323 gcaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 382
Db 1274 ggtccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1333
OY 383 gcaaggaagaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 442
Db 1334 gaaaggaagaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1393
OY 443 acttgcacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 502
Db 1394 atttcgacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1453
OY 503 tggccttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 562
Db 1454 tggcattctcgggagattgttataaacaccctgtttagcatgacgcaaatgtccagatga 1513
OY 563 gcttgaagagacatcagccgacacgagcgccgcatgtatccacatcggcgacaca 622
Db 1514 gggttaaagatatccacactgactgacggtgggagatgttaaccatattgacagaaga 1573
OY 623 agaccctgtgtgacacccgacggtgaggaagagccctgtgacgttggcgctgacgaatcgg 682
Db 1574 aaacgtgtgttagcacccgacaggtgttagaagaagcacttaagccttgggggtaactaaatcgg 1633
OY 683 tggagcgtcgtgacatcagcgtgtgagcgctgtgacgacgaccccaacacatacctgttctgc 742

```


QY	Db
23	2
24	2
25	2
26	2
27	2
28	2
29	2
30	2
31	2
32	2
33	2
34	2
35	2
36	2
37	2
38	2
39	2
40	2
41	2
42	2
43	2
44	2
45	2
46	2
47	2
48	2
49	2
50	2
51	2
52	2
53	2
54	2
55	2
56	2
57	2
58	2
59	2
60	2
61	2
62	2
63	2
64	2
65	2
66	2
67	2
68	2
69	2
70	2
71	2
72	2
73	2
74	2
75	2
76	2
77	2
78	2
79	2
80	2
81	2
82	2
83	2
84	2
85	2
86	2
87	2
88	2
89	2
90	2
91	2
92	2
93	2
94	2
95	2
96	2
97	2
98	2
99	2
100	2

[illegible]

```

; NUMBER OF SEQ ID NOS: 11
;
; SOFTWARE: PatentIn Ver. 2.1

```



```

; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,609
; FILING DATE: 23-Oct-1996
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 397-8338
; TELEFAX: (415) 397-8338
;
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-735-609-11

Query Match      23.0%; Score 241.2; DB 2; Length 699;
Best Local Similarity 77.0%; Pred. No. 1.6e-33;
Matches 294; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1 atgcccaagaagaaggaagtgagcaactgtctgacctgtgacccgagcaactgtccggcc 60
DB 318 atgcccaagaagaaggaagtgagcaactgtctgacctgtgacccgagcaactgtccgca 377
QY 61 ctgcccgtgagcgcacccagcgcagagtgcgcaagaacctgagtgacatgtttccgcgac 120
DB 378 ttaccggctgcatgcaacgagtgatgaggttccgaacacctgcatgacatgtttcagggat 437
QY 121 cgcacagccttcagcagcagcaacctgtgaagatgtctgtgagcgtgtgcccagctgggac 180
DB 438 cgcacagcctgtttctgacgacatcctggaataatgcttctgtccgtttgccggctggcgcg 497
QY 181 gctgtgtcaagctgaacaacgcgaagtgttcccgccgcgcgcgcgcgcgcgcgcgcgcgcgc 240
DB 498 gcattgctgcaagtggaatgacggaatggtttccgcgaacacctggaatggtttccggcat 557
QY 241 tactgtctgtacctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300
DB 558 tatcttctatattcttcagggcgccggtctgagcagtaaaactatccagcaaatgtttggcgc 617
QY 301 cagctgaacatgtctgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
DB 618 cagctaaacatgcttcatctgcgtccggcgccacgaccaagtgacgaatgctgtt 677
QY 361 agcctgtgtatgcccgcacatc 382
DB 678 tcactgtgtatgcccgcgcatcc 699

RESULT 9
US-09-315-372-11
; Sequence 11, Application US/09315372
```

```

; Patent No. 6057158
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Amalfitano, Andrea
; APPLICANT: Hauser, Michael A.
; APPLICANT: Kumar-Singh, Rajendra
; APPLICANT: Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,372
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/735,609
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
;
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-315-372-11

Query Match      23.0%; Score 241.2; DB 3; Length 699;
Best Local Similarity 77.0%; Pred. No. 1.6e-33;
Matches 294; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1 atgcccaagaagaaggaagtgagcaactgtctgacctgtgacccgagcaactgtccggcc 60
DB 318 atgcccaagaagaaggaagtgagcaactgtctgacctgtgacccgagcaactgtccgca 377
QY 61 ctgcccgtgagcgcacccagcgcagagtgcgcaagaacctgagtgacatgtttccgcgac 120
DB 378 ttaccggctgcatgcaacgagtgatgaggttccgaacacctgcatgacatgtttcagggat 437
QY 121 cgcacagccttcagcagcagcaacctgtgaagatgtctgtgagcgtgtgcccagctgggac 180
DB 438 cgcacagcctgtttctgacgacatcctggaataatgcttctgtccgtttgccggctggcgcg 497
QY 181 gctgtgtcaagctgaacaacgcgaagtgttcccgccgcgcgcgcgcgcgcgcgcgcgcgcgc 240
DB 498 gcattgctgcaagtggaatgacggaatggtttccgcgaacacctggaatggtttccggcat 557
QY 241 tactgtctgtacctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300
DB 558 tatcttctatattcttcagggcgccggtctgagcagtaaaactatccagcaaatgtttggcgc 617
QY 301 cagctgaacatgtctgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
DB 618 cagctaaacatgcttcatctgcgtccggcgccacgaccaagtgacgaatgctgtt 677
```

QY 181 g c c t g t g c a a g c t g a a c a c c g c a a g t g t t c c c c g c g a g c c g a g a c t g c g g a c 240

Oy 1 atgcccaagaagaagaaagttagcaacctctgacgcgtgcaccagaacctgcccgc 60
|||||
Db 318 ATGCCCAAGAGAAGAGGAGTGTCCAAATTTACTGCACGCTACACCACAAAATTGGCCGCA 377

RESULT 13
US-08-459-448A-6
; Sequence 6, Application US/08459448A
; Patent No. 5859336
; GENERAL INFORMATION:

Query Match	8.48;	Score 88.2;	DB 2;	Length 3624;
Best Local Similarity	47.0%;	Pred. No. 1.9e-07;		
Matches 310;	Conservative 0;	Mismatches 343;	Indels 6;	Gaps

RESULT 14
 US-08-459-595A-6
 ; Sequence 6, Application US/08459595A
 ; Patent No. 6018104
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Koziel, Michael G.
 ; APPLICANT: Desai, Nalini M.
 ; APPLICANT: Lewis, Kelly S.
 ; APPLICANT: Kramer, Vance C.
 ; APPLICANT: Warren, Gregory W.
 ; APPLICANT: Evola, Stephen V.
 ; APPLICANT: Crossland, Lyle D.
 ; APPLICANT: Wright, Martha S.
 ; APPLICANT: Merlin, Ellis J.
 ; APPLICANT: Lannis, Karen L.
 ; APPLICANT: Rochstein, Steven J.
 ; APPLICANT: Bowman, Cindy G.
 ; APPLICANT: Dawson, John L.
 ; APPLICANT: Dunder, Erik M.
 ; APPLICANT: Pace, Gary M.
 ; APPLICANT: Suttie, Janet L.
 ; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
 ; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
 ; NUMBER OF SEQUENCES: 94
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NO. 6018104artis Corporation
 ; STREET: Patent & Trademark Dept., 520 White Plains
 ; STREET: Rd., POB 2005
 ; CITY: Tarrytown
 ; STATE: New York


```

: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 3624 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: other nucleic acid
:   DESCRIPTION: /desc = "Synthetic DNA"
:   HYPOTHEICAL: NO
: FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 1..3621
:   OTHER INFORMATION: /product= "Full-length, maize
:   OTHER INFORMATION: optimized cry1B"
:   OTHER INFORMATION: /note= "Disclosed in Figure 6."
:
US-08-459-504B-6
```

```

Query Match      8.4%; Score 88.2; DB 3; Length 3624;
Best Local Similarity 47.0%; Pred. No. 1.9e-07;
Matches 310; Conservative 0; Mismatches 343; Indels 6; Gaps 1;

QY 347 acagaacgcgtgagcctgtgagtcgacgacacgcaagagagacgtgagccgcg 406
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 476 ACGCCAGCGCGCCACCTGCACTGCTGCTGCGCGACGCGCAGCTGTGCGCAGCG 535

QY 407 agcgcgcgaagcagccctgctgctgagcgcaacgcttcgacacaggtgcagcctga 466
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 536 AGTTGGGCTTACGACGAGAGATCCAGCGCTACTACGAGCGCCAGGTGAGAGCCACC 595

QY 467 tggagaacgcgacccgtgcgacgacatccgcaacgtgcttcttggcactgcctaca 526
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 596 GCGACTACGACGACTGACGCTGAGGTGATACACACCGGCTGACACGCTGCGCGCA 655

QY 527 acaccctctgctgcatcgcgagatcgccgcgcatccggtgaaagacatcagccgcg 586
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 656 CCAACGCGCGCCAGCTGGGTGCGCTACACAGTTCCGCGCGACTGACCTGGGCGTGC 715

QY 587 acgagcgccgcaatgtgataccacatcgccgcaacgaacccctgtgagcagcgcgcg 646
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 716 TGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775

QY 647 tggagaagccctgagcctgagcgctgacaaagctggtgagcgctgagatcagcgtag 706
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 776 CCCAG-----CTGACCGCGCGAGGTGTACACCGAGCCATCGGCGCACCGCGGTGACA 829

QY 707 ggtgtgcccgcagcccaacacactcgttctgctgctgctgcaagaagcgctgagcg 766
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 830 TGGCCAGCATGTGATGTGATACACACACACGCCCCCAGCTTCAGCGCCATCGAGGCGCG 889

QY 767 cccccagcgccacagcagcagcgtgagcaccgagccctgagaggaatctcgagccacc 826
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 890 CCATTCGCGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 949

QY 827 accgctgtatctacgagcgcaagagcagcagcgccacgtacacgtgctgagcgagcc 886
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 950 GCGGCTGGAGCAACACCCGCAATGACCTTATGCGCGCCACACCATCCAGAGCCGCC 1009

QY 887 acagcgccgcgctgagcgccgcgacatgacgacgagcgcgcgctgagcatcccgaga 946
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 1010 CCATTCGCGCGCGCTGAGACACACGACACCGGCGCCACACACGATCAACCCCG 1069

QY 947 tcatgagcgcgagcgctgagcaacgltgaaacatgltgataactaactacgcaactg 1005
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DY 1070 TGACCTGTGCTTCCGCGACCGGAGCTGTACCGCACCGAGGCTACGCGCGCGCTGTG 1128
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2002, 00:48:43 ; Search time 1779.62 Seconds
(without alignments)
7963.387 Million cell updates/sec

Title: US-09-662-128A-1
Sequence: 1 atgcccaagaagaagagagaa.....gcctcgtgagagcgcgcac 1050

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gq_estli:*
10: gq_estlc:*
11: gq_hc:*
12: gq_gss:*
13: em_gss_hum:*
14: em_gss_liv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	390	37.1	688	10	B1651236 603298328
2	231	22.0	431	9	AW211323
3	181.2	17.3	563	10	B1650191
4	158.8	15.1	340	9	AW212478
5	150.6	14.3	338	9	AW211972
6	128.2	12.2	343	9	AW210918
7	86	8.2	836	10	B1556329
8	79.8	7.6	288	12	FR0013439
9	75.6	7.2	925	12	CNS0091P
10	74.4	7.1	852	10	B1949774
11	74.4	7.1	854	10	BE261766
12	72.2	6.9	1464	3	BE636746
13	72	6.9	949	12	CNS031R8
14	69.8	6.6	616	12	CNS03VH
15	69	6.6	935	12	CNS006X
16	68.2	6.5	637	10	BE251052
17	67	6.4	692	10	BM134842

C	18	67	6.4	759	10	BF662500
	19	66.8	6.4	559	12	AZ661219
	20	66.8	6.4	610	10	BM324963
	21	66.8	6.4	1183	10	BE636683
	22	66.6	6.3	654	6	B140552
C	23	65.8	6.3	622	9	AV927327
	24	65.8	6.3	683	10	B1546421
	25	65.6	6.2	702	10	B1836679
	26	65.4	6.2	1101	12	CNS017SY
	27	65.2	6.2	1177	10	BE636682
	28	65	6.2	760	10	BE66871
	29	65	6.2	1453	10	BM320858
	30	64.8	6.2	503	10	BE404787
	31	64.8	6.2	964	10	BE6343549
	32	64.6	6.2	607	10	BE639506
	33	64.6	6.2	1239	12	AG047572
	34	64.4	6.1	591	10	B1718551
	35	64.4	6.1	857	10	BP262005
	36	64.2	6.1	485	9	AW161386
	37	64.2	6.1	674	10	B1955291
	38	63.8	6.1	824	10	B1953697
	39	63.4	6.0	598	9	AW324623
	40	63.4	6.0	710	9	AU163439
	41	63.4	6.0	819	10	BE309572
	42	63.2	6.0	598	10	BE518410
	43	63.2	6.0	632	10	B1959076
C	44	63.2	6.0	1137	10	BE809979
	45	63.2	6.0	1452	12	AG032979

ALIGNMENTS

RESULT 1
B1651236
LOCUS 603298328F1 NIH_CGAP_Mam3 Mus musculus CDNA clone IMAGE:5338806 5',
DEFINITION mRNA sequence.
ACCESSION B1651236
VERSION B1651236.1 GI:15565472
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaapb-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1860 row: h column: 07
High quality sequence stop: 688.
Location/Qualifiers
1. 688
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5338806"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPOrt6; site:1; Notf:
site:2; Salt: Cloned unidirectionally. Primer: Oligo dt.
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigators providing
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference

FEATURES

source

for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI-CGAP Library."

BASE COUNT 180 a 158 c 177 g 173 t

ORIGIN

Query Match 37.1%; Score 390; DB 10; Length 688;
Best Local Similarity 74.8%; Pred. NO. 5.1e-42;
Matches 489; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

```

OY 14 aagaagaagttagcaacctgtgacacgtgacacgaacctgcccctgcccgtgagc 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 AGTGTGAATTCATTTACTGACCGTACACAAAATTTGCTGCATTTACGGTGCATG 94
OY 74 ccaccacgacgaggtgtgcgaagaacctgtatgtacatgttccgaccgcgaagcttca 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 CAACGAGTATGAGTTCGCAAGAACCTGATGACATGTTACGAGATCGCAGCGGTTT 154
OY 134 gcgagcacacttgaaagtgtctgtgacgctgtgcccgaagctggcgccgtgtgcaagc 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 CTGACATACCTGGAAATGCTTCTGTCCTTTGCGGTCGTGGCGGATGTCAGT 214
OY 194 tgaacaacgcaagtgtgtcccgccgacgcccgaagacgtgtgcgaactactgtctaac 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 TGAATTAACCGGAATGTTTCCCGAGAACCTGAAGATGTTCGGATTAATCTTATATC 274
OY 254 tgcagcccgcgctgtgctgtgaaacacatccagcagcactgtggcgagctgagacatgc 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 TTTCAGGCGCGGCTGTGCGACTAAATACTATCCACAACTTTGGGCGACTAAACATGC 334
OY* 314 tgcacgcgcgacgagcgtctgcccgcacacgaagaacgcccgtgagcctgtgtatgc 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 TTTCATGTCGTGTCGGGCTGCACAGCAAGTACAGCAATGCTTTCACTGGTTATGC 394
OY 374 ggcgcacccgaagaagaacgtgtgacgcccgcgcgcgcgcgcgaagagcctgtgcttcg 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 GCGCGATCCGAAAGAAACGTTGATGCCGGAACGTCGCAAAACAGGCTCTAGCGCTTCG 454
OY 434 agcgacgcaacttgacacaggtgtgcgacgctgtatgagaacacgacgcctgtgcagaaca 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 AAGGCACTGATTTCCACACAGGTTCTTTCATCATATGAAATATGCGATGGCCAGATTA 514
OY 494 tccgcaacctgtgcttctgtggacatgcctacacaacctgtgtgcacatgcagatgc 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 TACGTAATCTGGCATTTCTGGGGATTTGCTTATTAACACCTGTTAGTACGCAAAATGG 574
OY 554 cccgacatccgctgtgagacatccagccgacgacgagcgcgcgcgcgcgcgcgcgcgcgc 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 CCAAGATCAAGGGTTAAAGTATCTACACTGACTGACGCGTGGAGATGTTAATCCATATTG 634
OY 614 gccgcaccaaagacctgtgtgacacccgcgcgcgtgtgagaagccctgtgacctgc 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 GCAGAACGAAACGCTGTTAGCACCGCAGGTGTAAGAGAACGACTTACCTGCG 688

```

RESULT 2
AW211323 431 bp mRNA linear EST 03-DEC-1999
LOCUS u079f01.y1 NCI-CGAP_Mam3 Mus musculus cDNA clone IMAGE:2648761 5'
DEFINITION similar to SW:RECR_BPPI P06956 RECOMBINASE CRE. [1] ;, mRNA
sequence.

ACCESSION AW211323
VERSION AW211323.1 GI:6517271
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE 1 (bases 1 to 431)
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

COMMENT

Other_ESTS: u079f01.x1
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
www.bio.lnlnl.gov/db/rrp/image/image.html

MG1:1029213
Seq primer: -40RP from Glibco
High quality sequence stop: 279.
Location/Qualifiers

FEATURES

1..431
/organism="Mus musculus"
/strain="129 - C57/B6 - FVB"
/db_xref="taxon:10090"
/clone_image="IMAGE:2648761"
/clone_lib="NCI-CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="PH10B"
/note="Organ: mammary; Vector: PCMY-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

BASE COUNT 118 a 101 c 111 g 101 t

ORIGIN

Query Match 22.0%; Score 231; DB 9; Length 431;
Best Local Similarity 71.0%; Pred. No. 2.3e-21;
Matches 306; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```

OY 197 acaacgcgaagtgtgtcccgccgacgacgagacgtgtgcagcactcctgtgttacctgc 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATAAACCGAATAGTTTCCCGCAGAACCTGAAGATGTTCCGATTTATCTTATATCTTC 60
OY 257 aggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AGCGCGCGGCTGTGCGAGTAAATACTATCCAGCAACTTTGGGCCAGCTAAACATGCTTC 120
OY 317 accgcgcgaagcgcgtgtcccgcccgacgacgacgacgacgacgacgacgacgacgacgacg 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 ATCGTCGGTCCGGGCTGCCACAGCAAGTACAGCAATGCTGTTTCACTGCTTATGCGGC 180
OY 377 gcatcgcgaagagaacgtgtgacgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GGATCCGAAAGAAAGAAAGCTTATGATCCCGGTGACCTGCAAAACAGGCTCTAGCGTTCGAC 240
OY 437 gcaccgaacttcgacacaggtgtgcgacgctgtatgagaacacgacgcgcgtcccaagacatcc 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GCACGTGATTTGCACAGAGTTGCTTCACTCATGCAAAATAGAGATCGCTGCCAGATATATAC 300
OY 497 gcaactgtgaccttctgtgacatgcctacaaacacctgtgcgcgcgcgcgcgcgcgcgcgcgc 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GTAACTGCGCACTTCTGGGATTTGCTTATTAACACCCGGTACCTATATAGCCGAAATGCCA 360
OY 557 gcatcgcgttgaagacatccgacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GGATCAGGTTAAAGATATCTCACAGGCTCAGCGTGGAGAAATGTAATCATATTGGGA 420
OY 617 gaaccaagacc 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 CAACGGAAGACC 431

```

RESULT 3
B1650191

LOCUS	B1650191	563 bp	mRNA	linear	EST 12-SEP-2001
DEFINITION	6032966283p1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5336853 5',				
ACCESSION	mRNA sequence.				
VERSION	B1650191				
KEYWORDS	B1650191.1 GI:15564427				
SOURCE	EST.				
ORGANISM	house mouse.				
	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 563)				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

FEATURES
 Plate: L1AM11855 row: f column: 22
 High quality sequence stop: 563.
 Location/Qualifiers

source

1.563

1

MG1:10302

164

1

BASE COUNT
ORIGIN
143 a 133 c 149 g 138 t
Site 2: Organ: Mammary; Vector: pCMV-SpRt6; Site: 1; NCI.
Average insert: Cloned unidirectionally. Primer: Oligo dt.
Technologies: catalog#227/18. Invested by Life
samples: Lohrer Hemiploidy/Chu Xia Deng, NIH Reference
(1999). Note: This is a NCI-CCGAT Library.

Query Match	17.3%	Score 181.2;	DB 10;	Length 563;
Best Local Similarity	76.6%	Pred. No. 6.5e-15;		
Matches 222; Conservative	0;	Mismatches 68;	Indels 0;	Gaps 0

BASE
ORIG

COUNT
IN

75 a

22, 37-43
100 C

(1999). "79 g

86 f

[illegible]

RESULT	4
AM212478/c	
LOCUS	
DEFINITION	
AM212478	
u089c09.x1	
NCI_CGAP_Mam3	
Mus musculus	
muscle	
cdna	
clone	
IMAGE:2649712	
EST_03-DEC-1999	
linear	
mRNA	
340 bp	

Similar to SW:RECR_BPPI P06956 RECOMBINASE CRE. [1] : mRNA
sequence.
AW212478
NM212478.1 GI:6518565
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 340)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: u089c09.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nlm.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
DNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbfp/image/image.html

BASE COUNT	75 a	100 c	79 g	86 t
ORIGIN				

Query Match	15.1%;	Score 158.8;	DB 9;	Length 340;
Best Local Similarity	76.0%;	Pred. No. 5.4e-12;		
Matches 196;	Conservative 0;	Mismatches 62;	Indels 0;	Gaps 0;

QY	192	gaccgggcccctggagggacatcttcgaggcgacacccgcgtatctcaagcgccgaaga	851
Db	293	CAATCGGCGCTGGAGAGGATTTTGGAGCACTCATCATTTATTTACGGCCCTTAGGA	234
QY	852	cgaaacgcggccagcgcgtacctgycgtgagcggccacagcgcgcgcgtlggcgccgcgcg	911
Db	233	TGATTTCTGGTACGAGATACCTGGCTGGTCTGGACACAGTGGCCGTTGGAGAGCCGGG	174
QY	912	cgaaatggcccgccgcggcggtgagcatcccccagatcatgacagcgccgcgttgaacaa	971
Db	173	AGATATGGCCCGGCTGGAGTCTTTCATATACCGGAGATCATGCAAGCTGTGGCTGGACCA	114
QY	972	cgltgaacatcgtatgagacacatccgcgaacccctggagcagcggagccgcgcgatitggg	1033
Db	113	TGTAAATATTGTATGACATATATCCGTAACTCGGATAGTAGAACAGGGGCAATGATGGTG	54

QY	1032	ccctgctgagagcgcga	1049
Db	53	CCTGCTGAAGATGGCGA	36

```
RESULT 5
AM211972/c 338 bp mRNA linear EST 03-DEC-1999
LOCUS uc079f01.x1 NCI CGAP Mam3 Mus musculus cDNA clone IMAGE:2648761.3
DEFINITION Similar to SW:RECR_BPPI P06956 RECOMBINASE CRE. [1] ;, mRNA
sequence.
ACCESSION AM211972
VERSION AM211972.1 GI:6517961
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogamathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 338)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: uc079f01.y1
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
MGI:1029213
Seq primer: -40UP from GIBCO
High quality sequence stop: 151.
FEATURES
Source
Location/Qualifiers
1..338
/organism="Mus musculus"
/strain="129 - C57/B6 - FVB"
/db_xref="taxon:10090"
/clone="IMAGE:2648761"
/clone_1ib="NCI CGAP Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT 78 a 93 c 82 g 85 t
ORIGIN
Query Match 14.3% Score 150.6; DB 9; Length 338;
Best Local Similarity 76.7%; Pred No. 6.3e-11;
Matches 197; Conservative 0; Mismatches 59; Indels 1; Gaps 1;
QY 793 accgggacctgagggcattctcgagggccaccacgcctgatacggcgcaagagc 852
Db 292 ACTGCGCCCTGGAGAGGAT-TTCTAAGCCACATCATGATTGATTAACGCGCTAAGGAT 234
QY 853 gaacggcgcaagcgctactgctgtagagcgccacagcgcggtggcgccgcgcg 912
Db 233 GACTTGGTGAAGATACCTGGCTGTCTGGACACAGTCCCGTCTCGAGCGGCTTA 174
QY 913 gacatggcccgcgcgctgagcattcccgagatcatcgagcgcggtgtagaacac 972
Db 173 CATATGCGCCGCGCTGAGATTTCATATACCGAGATCATGAACTGTGTGGCTGACCAAT 114
QY 973 gtgaacactgtagaatacaccgaactggaacacggcgagagcgcgcatggtgcgc 1032
Db 113 GTAATATTTGTCAATGAACTATATCCGTAACGTGATAGTGAACAGGGGCAATGTGCGC 54
QY 1033 ctctgtgagagcgcgca 1049
```

```
Db 53 CTGCTGGAAGATGGCGCA 37
||||||| || |||||
RESULT 6
AM210918 343 bp mRNA linear EST 03-DEC-1999
LOCUS uc089c09.y1 NCI CGAP Mam3 Mus musculus cDNA clone IMAGE:2649712.5
DEFINITION Similar to SW:RECR_BPPI P06956 RECOMBINASE CRE. [1] ;, mRNA
sequence.
ACCESSION AM210918
VERSION AM210918
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogamathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 343)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: uc089c09.x1
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
MGI:1030164
Seq primer: -40RP from GIBCO
High quality sequence stop: 222.
FEATURES
Source
Location/Qualifiers
1..343
/organism="Mus musculus"
/strain="129 - C57/B6 - FVB"
/db_xref="taxon:10090"
/clone="IMAGE:2649712"
/clone_1ib="NCI CGAP Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT 95 a 97 c 85 g 84 t 2 others
ORIGIN
Query Match 12.2% Score 128.2; DB 9; Length 343;
Best Local Similarity 66.5%; Pred No. 5.1e-08;
Matches 228; Conservative 0; Mismatches 110; Indels 5; Gaps 3;
QY 319 cggcgaacggcgctgcccgcggccgacgacgacgacgctgtagctgtagtcgcgcg 378
Db 2 CGTGGTCCGGGCTGCGCCGACGACCAAGTACAGCAATCTCTTACCTGATTATCCGCGG 61
QY 379 atccgaaggaagacgtgtagcgccggcgagcgccgaagcgccgcttggagcgc 438
Db 62 ATCCAAAGAAAGAGGTTGATCCGGTGAACGTCGCAAAAGGCTCTAGCCTTGAACGC 121
QY 439 accgagcttcgaacgaagctgtagcgccgtagtagaagacgacgacgctgcacgacatccgc 498
Db 122 ACTGATTTTC-ACCAAGTTCGTTCACTCATGGAATAACGATCCGTCGACAGATATACGT 180
QY 499 aacctgcttctcttgagca-tgcctacaacacctgctgcgcattgcgcgagatgcgcgc 557
```

RESULT	8	FR0013439	288 bp	DNA	linear	GSS 18-SEP-199
LOCUS		F.rubripes GSS sequence, clone 122014bG11,				genomic survey sequence
DEFINITION		AL004689				
ACCESSION		AL004689.1	GI:2450259			
VERSION		GSS: genome survey sequence.				
KEYWORDS		Takifugu rubripes.				
SOURCE		Takifugu rubripes.				
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;				
		Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;				
		Tetraodontidae; Takifugu.				
		1 (bases 1 to 288)				
REFERENCE		Elger,G., Clark,M., Smith,S., Meek,S., Warner,S., Umranta,Y.,				
AUTHORS		Williams,G. and Brenner,S.				
TITLE		Direct Submission				
JOURNAL		Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource				
COMMENT		Centre Hinxton, Cambridge, CB10 1SB. Email: blchelp@hgm.mrc.ac.uk				
		Vector: pBluescript II KS				
		Y_type: phagemid				
		PRIMER: KS				
		DESCR:				
		One pass dye-terminator sequencing of cosmid cloned genomic				
		sequence.				
FEATURES		Location/Qualifiers				
source		1..288				
		/organism="Takifugu rubripes"				
		/db_xref="taxon:31033"				
		/clone_11b="cosmid 122014"				
		/clone="122014bG11"				
BASE COUNT		72 a 70 c 70 g 73 t				3 others
ORIGIN						
Query Match		7.6%;	Score 79.8;	DB 12;	Length 288;	
Best Local Similarity		80.9%;	Pred. No. 0.099;			
Matches		93;	Conservative 0;	Mismatches 22;	Indels 0;	Gaps 0.
Oy	1	atgcccaagaagaagaagtgcgaacctctgcacgctgcaccagaacctgccgcgc	60			
Db	54	atgcccaagaagaagaagtgcgaacctctgcacgctgcaccagaacctgccgcgc	113			
Oy	61	ctgcccgctgcgaaccgacgacgaagtgcgcaagaacctgatactgttc	115			
Db	114	ttaccgctgatgatgcacgacgctgatgaggttcgcaagaacctgatgcacatgttc	168			
RESULT	9	CNS0091P	925 bp	DNA	linear	GSS 03-JUN-1999
LOCUS		Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION		BAK119D16 of RPI-98 library from Drosophila melanogaster (fruit				
		fly), genomic survey sequence.				
ACCESSION		AL053013				
VERSION		AL053013.1	GI:4934461			
KEYWORDS		GSS.				
SOURCE		fruit fly.				
ORGANISM		Drosophila melanogaster				
		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
		1 (bases 1 to 925)				
REFERENCE		Genoscope.				
AUTHORS		Direct Submission				
TITLE		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;				
JOURNAL		Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT		- web : www.genoscope.cns.fr)				
		Determination of this BAC-end sequence was carried out as part of a				
		collaboration with the Berkeley Drosophila Genome Project (BDGP).				
		The BDGP is constructing a physical map of the genome of the fruit fly				

Db 237 CTTGCGGCGCATGCGCACTCAACGCGGCGCTTCCGCGGCGGCTCGCGCCATCA 296
 QY 819 ggcacccacccgctgtatcatcagcgccaaagagacagcgccgagctactgctg 878
 Db 297 GGTGTCCACCTTCATCCGTTTGACAGAGTGGCCAACTCTTCCGCCCTCCGCGGCGT 356
 QY 879 gagcgccacagcgcccgctggtggcgccgacacatgcccgcgcgctgagcat 938
 Db 357 GTTCAAGAACGCTTACATGACGAGTGCGCCGCTCTCGGAGACACCGCGGCCCT 416
 QY 939 ccccgagatcatcgagcgcgcgctgtgacacagctgaacatcgatgaactacatcg 998
 Db 417 GGTGCGCAACGCTTACCTTACTTGGCGTACGACACCCCGGAGCATCAGCCTGAA 476
 QY 999 caactgtgacagcgagacggcgccatggtgctgctgtgagagcg 1046
 Db 477 CTACCGCAGCTTCCAGCGGCGGACCAACCGTGGTGTACCAAGAACACGG 524

RESULT 11
 LOCUS BF261766 854 bp mRNA linear EST 23-OCT-2001
 DEFINITION HV.CEA0002C17f Hordeum vulgare seedling green leaf EST library
 HVCDNA0004 (Blumeria challenged) Hordeum vulgare cDNA clone
 BF261766
 ACCESSION BF261766 GI:13259943
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Hordeum vulgare
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Poideae
 : Triticeae: Hordeum.
 1 (bases 1 to 854)

AUTHORS
 Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Wei, F., Begum, D.,
 Fritsch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi,
 D.W., Penton, R.D., Oates, R. and Main, D.
 Development of a genetically and physically anchored EST resource
 for barley genomics: Blumeria infected incompatible (Mla13)
 seedling leaf cDNA library

JOURNAL
 COMMENT
 On Nov 17, 2000 this sequence version replaced gi:11192753.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall
 Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 509
 Seq primer: AATTACCTCACTAAGG
 High quality sequence stop: 727.
 Location/Qualifiers

FEATURES

source

1..854
 /organism="Hordeum vulgare"
 /cultivar="C116155 (Mla13)"
 /db_xref="taxon:4513"
 /clone="HV.CEA0002C17f"
 library HVCDNA0004 (Blumeria challenged)
 /tisue="seedling green leaf"
 /lab_host="TJ121"
 /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
 C.I. 16155 (Mla13) plants were greenhouse grown in the R
 wise lab at Iowa State University, Ames, IA; 7 day old
 green seedlings were challenged with isolate A27 (AVRMla13
) of Blumeria graminis f. sp. hordei, and leaves were
 harvested 20 and 24 hr post-inoculation and snap frozen;
 uninoculated leaves were harvested 20 hr post-inoculation
 (Wei, Wise). In the TJ Close lab at the University of
 California, Riverside, total RNA was prepared from each
 sample pool, equal quantities of all three RNA pools were
 combined, poly(A) RNA was purified from the mixture, one

cDNA library was made, and 1 million pfu were in vivo
 excised to give pBluescript SK(-) cDNA phagemids (Choi,
 Close). Phagemids were plated and picked at the Clemson
 University Genomics Institute (CUGI) (Begum, Palmer,
 Fritsch, Atkins and Wing). Plasmid DNA preparations, DNA
 sequencing and sequence analysis were performed at CUGI
 (Wing, Yu, Fritsch, Henry, Simmons, Oates, Rambo, Main).
 The sequence has been trimmed to remove vector sequence
 and contains a minimum of 100 bases of phred value 20 or
 above. For more details on library preparation and
 sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
 this clone see <http://www.genome.clemson.edu/orders>
 see Close et al., Wing, R., Kleinhofs, A., Wise, R. (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (<http://wheat.pw.usda.gov/gnpages/bgrn/31/cover.html>)"

BASE COUNT 151 a 304 c 249 g 149 t 1 others

Query Match 7.1%; Score 74.4; DB 10; Length 854;
 Best Local Similarity 47.4%; Pred. No. 0.45;
 Matches 222; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

QY 579 ccgacacagcgcgccgctgtatccatcgccgacacgaacacctgtgagcac 638
 Db 258 CCGCACTCCGCAATCGGCTCATCTCTGACACGGCAAGACCACTCGCCCAATCGC 317
 QY 639 cgcgcgctggaagaagccctgagcctggtgagcgagcaagctgtgagcgtgac 698
 Db 318 CGCAGACCTCCACAGCGGCGCTCTGGGTCAGAACACGTCGCGCTTACCTTGC 377
 QY 699 cgtgagcgcgctgtgacgagaccccaactactgttctgcgctgtgcaagaacg 758
 Db 378 CGTAACATCAATCAATGCGCCCGCAACGAGGTGATGGCGGCGCACAGAGAT 437
 QY 759 cgtgacgcccccgacgacgacgacgacgacgacgacgacgacgacgacgac 818
 Db 438 CTTGCGGCGCATGCGCACTCAACCGGCGCTCTCGGCGGCGGCGGCGCATCA 497
 QY 819 ggcacccacccgctgtatcatcagcgccaaagagacagcgccgagctactgctg 878
 Db 498 GGTGTCCACCTTCATCCGTTTGACAGAGTGGCCAACTCTTCCGCCCTCCGCGGCGT 557
 QY 879 gagcgccacagcgcccgctgtggtggcgccgacatgtgcccgcgcgtgagcat 938
 Db 558 GTTCAAGAACGCTTACATGACGAGTGCGCCGCTCTCGGAGACACCGCGCGCT 617
 QY 939 ccccgagatcatcgagcgcgcgctgtgacacagctgaacatcgatgaactacatcg 998
 Db 618 GGTGCGCAACGCTTACCTTACTTGGCGTACGACACCCCGGAGCATCAGCCTGAA 677
 QY 999 caactgtgacagcgagacggcgccatggtgctgctgtgagagcg 1046
 Db 678 CTACCGCAGCTTCCAGCGGCGGACCAACCGTGGTGTACCAAGAACACGG 725

RESULT 12
 BE636746
 ID BE636746 standard; RNA; EST; 1464 bp.
 AC BE636746;
 XX
 SV BE636746.1
 XX
 DT 28-AUG-2000 (Rel. 64, Created)
 DT 28-AUG-2000 (Rel. 64, Last updated, Version 1)
 DE rockefeller.0.487 Mastigamoeba balamuthi lambda ZAP II Library Mastigamoeba
 DE balamuthi cDNA similar to alcohol dehydrogenase (Adh), mRNA sequence.
 XX
 KW EST.

Fri May 10 08:35:57 2002

us-09-662-128a-1.rst

Page 11


```

134  Iaglygluargalalysglinalaleualahegluargthrasppheasp 150
      |||
3175  CCGGTGAACGTGCAAAACAGCGCTAGCGTCGAAACGACGATTTTGAC 3224
      |||
151  GlnValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAs 167
      |||
3225  CAGGTTCCGTCACCTCATGGAATAAGCATCGCTCCAGGATATACCTAA 3274
      |||
167  nleualaheleuglyllealaTyraThrLeuLeuArgllealaglul 184
      |||
3275  TCTGGCATTTCTGGGATTCCTTATTAACCCCTGTACGTATACCGAAA 3324
      |||
184  lealaarglleargValylsaspIleSerArgThrAspGlyArgMet 200
      |||
3325  TTGCCCGAGTTCAGGGTTAAAGTATCTACGACTGACGGGGGAGATG 3374
      |||
201  leuilehisileglyArgThrThrLeuValSerThrAlaGlyValG1 217
      |||
3375  TTAATCCATATTCGCAAGAAACGCTGGTTACACCGAGGTGTAGA 3424
      |||
217  ulysAlaLeuSerLeuGlyValThrLysLeuValGluArgThrPileserv 234
      |||
3425  GAAGGCACTTAGCCTGGGGGTAACTGGTCGAGCGATGGATTTCCG 3474
      |||
234  alserglyValAlaAspAspProAsnAsnTyrlLeupheCysArgValArg 250
      |||
3475  TCTCTGTGTAGCTGATGATCCGAAATACCTACTGTTTGGCGGTGAGA 3524
      |||
251  LysAsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAl 267
      |||
3525  AAAAAGGAGTGTGCGCGCATCTGCGACACGACGATACCACTGCGGC 3574
      |||
267  aleuGluGlyllePheGluAlaThrHisArgLeuIleTyrglyAlaLysA 284
      |||
3575  CCTGGAGAGGATTTTGAACCAATCATGATGATTTACGGCGGTAAAG 3624
      |||
284  spAspSerGlyGlnArgTyrlLeuAlaTrpSerGlyHisSerAlaArgVal 300
      |||
3625  ATGACTGTGTGACAGATACCTGCTGCTGAGCACAGTGCCTGCTGC 3674
      |||
301  G1ValAlaAlaArgAspMetAlaArgAlaGlyValSerIleProGluIleWe 317
      |||
3675  GGAGCCCGCGAGATATGCGCGCTGGAGTTTCAATACCGAGATCAT 3724
      |||
317  tGlnAlaGlyIleTyrlThrAsnValAsnIleValMetAsnTyrlIleArgA 334
      |||
3725  GCAAGCTGGGTGGCGGACCAATGTAATATTGTCATGAACTATATCCGTA 3774
      |||
334  snLeuAspSerGluThrGlyAlaMetValArgLeuLeuGluAspGlyAsp 350
      |||
3775  ACCTGGATAGTGAACAGGGGCAATGCTGCTGCTGGAAGATGGCGAT 3824
      |||
seq_name: gb_pat:AX191663
seq_documentation_block:
LOCUS AX191663 4960 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 29 from Patent WO0149832.
ACCESSION AX191663
VERSION AX191663.1 GI:15209844
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 4960)
AUTHORS Schwenk, F.
TITLE Transduction of recombinases for inducible gene targeting
JOURNAL Patent: WO 0149832-A 29 12-JUL-2001;
ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES
Location/Qualifiers
source 1..4960
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pCMV-I-Cre-pA"

```

```

BASE COUNT 1225 a 1213 c 1296 g 1226 t
ORIGIN
alignment_scores:
Quality: 1797.00 Length: 350
Ratio: 5.134 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-662-128A-2 x AX191663
Align seg 1/1 to: AX191663 from: 1 to: 4960
1 MetProLysLysArgLysValSerAsnLeuLeuThrValHisGlnAs 17
|||||
977 ATGCCCAAGAACAGAGAGAGTGTCCAAATTACTGACCGTACCCAAAA 1026
|||||
17 nleuProAlaLeuProValAspAlaThrSerAspGluValArgLysAsnL 34
|||||
1027 TTGGCTGCATTACCGGTGTCGATGCAACGAGTATGAGGTTGCAAGAACC 1076
|||||
34 MetLeuAspMetPheArgAspArgGlnAlaPheSerGluHisThrTrpLys 50
|||||
1077 TGAATGACATGTTTCAGAGGATCCGACGGGTTTCTGACATACCTGGAAA 1126
|||||
51 MetLeuLeuSerValCysArgSerTrpAlaAlaTrpCysLysLeuAsnAs 67
|||||
1127 ATGCTTCTGTCCGTTTGGCGGTCTGCGGCAATAAACTATCCACCACTTTGGGC 1176
|||||
67 nArgLysTrpPheProAlaGluProGluAspValArgAspTyrlLeuLeu 84
|||||
1177 CCGGAATAGTGTCCCGCAACCTGAAGATGTTGGCATATCTTCAT 1226
|||||
84 yrlLeuGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGly 100
|||||
1227 ATCTTCAGGCGCGCGGTCTGCGAGTAAAACTATCCACCACTTTGGGC 1276
|||||
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSe 117
|||||
1277 CACTTAAACATGCTTCATGCTGCGTCCGGCTCCACACCAAGTACAG 1326
|||||
117 rAsnAlaValSerLeuValMetArgArgIleArgLysGluAsnValAspA 134
|||||
1327 CAAATGCTGTTCACCTGTTATGGCGGAGATCCGAAACGAAACGTTGATG 1376
|||||
134 Iaglygluargalalysglinalaleualahegluargthrasppheasp 150
|||||
1377 CCGGTGAACGTGCAAAACAGCGCTTAGCGTTGGAAGCATGATTTGCAC 1426
|||||
151 GlnValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAs 167
|||||
1427 CAGGTTCTGTCACCTCATGGAATAATACGATCGTGCAGAGATATACGTAA 1476
|||||
167 nleualaheleuglyllealaTyraThrLeuLeuArgllealaglul 184
|||||
1477 TCTGGCATTTCTGGGATTCCTTATTAACACCCGTTACTATTAAGCGGAAA 1526
|||||
184 lealaarglleargValylsaspIleSerArgThrAspGlyArgMet 200
|||||
1527 TTGCCAGATCAGGGTTAAAGATATCTCACGTACTACCGTGGGAGAAAG 1576
|||||
201 leuilehisileglyArgThrThrLeuValSerThrAlaGlyValG1 217
|||||
1577 TTAATCCATATTCGCAAGAAACGCTGGTTACACCGAGGTGTAGA 1626
|||||
217 ulysAlaLeuSerLeuGlyValThrLysLeuValGluArgThrPileserv 234
|||||
1627 GAAGGCACTTAGCCTGGGGGTAACTAAAGTGTGCGAGCGATGGATTTCCG 1676
|||||
234 alserglyValAlaAspAspProAsnAsnTyrlLeupheCysArgValArg 250
|||||
1677 TCTCTGTGTAGCTGATGATCCGAAATACCTACTGTTTGGCGGTGACAGA 1726

```

```
251 LysAsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAl 267
|||||
1727 AAAAATGGTGTGCGCCCATTCGCCACAGCAGCATATCAACTGCGGC 1776
267 aleuGluGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAlaLysA 284
|||||
1777 CCGGAAGGAGATTTTGAAGCAACTCATCATGATTGATTACGGCGCTAAG 1826
284 spAspSerGlyGlnArgTyrLeuAlaIleTrpSerGlyHisSerAlaArgVal 300
|||||
1827 ATACCTCTGTGTCAGAGATACCTGGCTGGTCGACACAGTCCGCTGTC 1876
301 GlyAlaAlaArgAspMetAlaArgAlaGlyValSerIleProGluIleMe 317
|||||
1877 GGAGCCGCGCAGATATGCGCCGCTGAGTTTCAATACCGGAGATCAT 1926
317 tGlnAlaGlyGlyTyrPThrAsnValAsnIleValMetAsnTyrIleArgA 334
|||||
1927 GCAAGCTGGTGGCTGGACCAATGTAAATATTGTCATGAATATATCCGTA 1976
334 snLeuAspSerGlnThrGlyAlaMetValArgLeuLeuGluAspGlyAsp 350
|||||
1977 ACCTGGATGTGAACAGGGCATGTGCGCTGCTGGAAAGATGGCGAT 2026
seq_name: gb_pat:AX114854
seq_documentation_block:
LOCUS AX114854 5365 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 14 from Patent WO0129208.
ACCESSION AX114854
VERSION AX114854.1 GI:14031796
KEYWORDS:
SOURCE:
ORGANISM: synthetic construct.
artificial sequence.
REFERENCE:
AUTHORS: Kuehn, R., von Melchner, H. and Altschmid, J.
TITLE: Conditional gene trapping construct for the disruption of genes
JOURNAL: Patent: WO 0129208-A 14 26-Apr-2001;
ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)
FEATURES
Source
1..5365
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector PBADp-pgkCre"
BASE COUNT 1208 a 1448 c 1402 g 1307 t
ORIGIN
alignment_scores:
Quality: 1797.00 Length: 350
Ratio: 5.134 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-662-128a-2 x AX114854 ..
Align seg 1/1 to: AX114854 from: 1 to: 5365
1 MetProLysIlyAsArgLysValSerAsnLeuThrValHisGlnAs 17
|||||
1972 ATGCCCAAGAGAAAGAGAGAGTGTCCATTACTACGCTACACCAAA 2021
17 nLeuProAlaLeuProValAspAlaThrSerAspGluValArgLysAsn 34
|||||
2022 TTTCCTGCTATACCGCTGATGCAACGAGTATGAGGTGCAAGAAC 2071
34 euMetAspMetPheArgAspArgGlnAlaPheSerGluHisThrTyrLys 50
|||||
2072 TGTATGACATGTTTCAGGGATCCCGAGCGCTTTTTCAGACATACCTG 2121
51 MetLeuSerValCysArgSerTrrpAlaIleTrrpCysLysLeuAsnAs 67
|||||
2122 ATGCTTCTGTGCGCTTGGCGGTCGTGGCGCATGTCGAATTTAAATTA 2171
67 nArgLysTrrpPheProAlaGluProGluAspValArgAspTyrLeuLeu 84
|||||
2172 CCGGAATGGTTTCCCGCAGAACTGAAGATGTGCGCATTTATCTTCTAT 2221
84 yrLeuGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGly 100
|||||
2222 ATCTTCAGCGCGCGCGGTCTGGCAGTAAAACTATCCACCAATTGGGC 2271
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSe 117
|||||
2272 CAGCTAAACATGCTTCATCTGCTGGTCCGGGCTCCAGACCAAGTACAG 2321
117 rAsnAlaValSerLeuValMetArgArgIleArgLysGluAsnValAspA 134
|||||
2322 CAATGCTGTTTCAGTGTATATGCGCGGATCCGAAACAAACGTTGATG 2371
134 IagLysGluArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheAsp 150
|||||
2372 CCGGTGAACGTGCMAAACAGGCTCTAGCGTTTCGAACGACATTTGAC 2421
151 GlnValArgSerLeuMetCylAsnSerAspArgCysGlnAspIleArgAs 167
|||||
2422 CAGGTTGCTTCATCTCATGTGAATAAGCATCGCTGCCAGATATACGTA 2471
167 nLeuAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGlu 184
|||||
2472 TCTGCAATTTCTGGGATGCTGTTTAAACACCTGTTACGTATAGCCGAA 2521
184 IeAlaArgIleArgValLysAspIleSerArgThrAspGlyArgMet 200
|||||
2522 TTGCAGAGATCAGGGTTAAAGATATCTACGTACTGACGCTGGAGAA 2571
201 LeuIleHisIleGlyArgThrLysThrLeuValSerThrAlaGlyValG 217
|||||
2572 TTATATCATATTTGGCAGAACGAAACGCTGTGTAGACCGCAGCTGTAA 2621
217 uLysAlaLeuSerLeuGlyValThrLysLeuValGluArgTrrpIleSerV 234
|||||
2622 GAAAGCAGCTTAGCCTCGGGGCTAACTAACTGTGTGAGCGCATGATTCCG 2671
234 aIleSerGlyValAlaAspAspProAsnAsnTyrIleuPheCysArgValArg 250
|||||
2672 TCTGTGCTGTAAGCTGATGATCCGAATPACTACTGTTTGGCGGTCAGA 2721
251 LysAsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAl 267
|||||
2722 AAAAATGCTGTGCCGCGCATCTGCCACACGACATCTCAACTCCGCC 2771
267 aleuGluGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAlaLysA 284
|||||
2772 CCTGGAAAGGAGATTTTGAAGCAACTCATCATGATTGATTACGGCGCTAAG 2821
284 spAspSerGlyGlnArgTyrLeuAlaIleTrpSerGlyHisSerAlaArgVal 300
|||||
2822 ATGACCTGTGTCAGAGATACCTGGCTGTGACACAGTCCCGCTGTC 2871
301 GlyAlaAlaArgAspMetAlaArgAlaGlyValSerIleProGluIleMe 317
|||||
2872 GGAGCCGCGCAGATATGCGCCGCTGGAGTTTCAATACCGGAGATCAT 2921
317 tGlnAlaGlyGlyTyrPThrAsnValAsnIleValMetAsnTyrIleArgA 334
|||||
2922 GCAAGCTGGTGGCTGGACCAATGTAAATATTGTCATGAATATATCCGTA 2971
334 snLeuAspSerGlnThrGlyAlaMetValArgLeuLeuGluAspGlyAsp 350
|||||
2972 ACCTGGATGTGAACAGGGCATGTGCGCTGCTGGAAAGATGGCGAT 3021
seq_name: gb_sy:AY056050
```

```

seq_documentation_block:
LOCUS      AY056050                1080 bp    DNA     linear     SYN 19-NOV-2001
DEFINITION Synthetic construct codon optimized cre recombinase (iCre) gene,
complete cds.
ACCESSION  AY056050
VERSION    AY056050.1 GI:17016299
KEYWORDS
SOURCE     synthetic construct.
           synthetic construct.
           artificial sequence.
REFERENCE  1 (bases 1 to 1080)
AUTHORS   Shimshak,D.R., Kim,J., Huebner,M.R., Spergel,D.J., Buchholz,F.,
           Stewart,F.A., Seeburg,P.H. and Sprengel,R.
TITLE      Codon-Optimized Cre Recombinase Expression in the Mouse
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1080)
AUTHORS   Shimshak,D.R., Kim,J., Huebner,M.R., Spergel,D.J., Buchholz,F.,
           Stewart,F.A., Seeburg,P.H. and Sprengel,R.
TITLE      Direct Submission
JOURNAL    Submitted (13-SEP-2001) Molecular Neurobiology,
           Max-Planck-Institute, Jahnstr. 29, Heidelberg 69120, Germany
FEATURES
Source     1..1080
           /organism="synthetic construct"
           /db_xref="taxon:32630"
           /note="derived from bacteriophage P1"
misc_feature 1..13
           /note="contains Kozak sequence"
           14..1069
           /gene="iCre"
           14..1069
           /gene="iCre"
           /note="contains part of the large T nuclear localization
           signal sequence"
           /codon_start=1
           /transl_table=11
           /product="codon-optimized cre recombinase"
           /protein_id="AAJ31698.1"
           /db_xref="GI:17016300"
           /translation="MWPKKRKRVSNLLTVHONIPALPDATSDVRRKNIMDMFRDRA
           FSEHTMKMLISVCRSNAAMCKLNNRKWFAPEPDVADYLLYLQAKGLAVKTIQHLGQ
           LNMJHRSGLPSPDSNAVASLVMRRIRKENVDAGEBAKQALAEIRDFQVSLIENS
           DRQDDINLAFLIAYNTLRIAEIARIKVDISRTDGGRLIHIGRTLVSTIGVE
           KALSIGTKIYERVISYSGVADDPNNYLFCRVKNGVAAPSAISOLSTALLEGIEAT
           HRLTYGAKDSDGQRIYLAWSHSAKRVGAARMAAGASIPETIQAGGWTNVTVMYIR
           NIDSETGAMVRLLEDGP"
BASE COUNT 248 a      294 c      319 g      219 t
ORIGIN
alignment_scores:
Quality: 1793.00      Length: 350
Ratio: 5.123          Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.714
alignment_block:
US-09-662-128a-2 x AY056050
Align seg 1/1 to: AY056050 from: 1 to: 1080
1 MetProLysLysLysArgLysValSerAsnLeuLeuThrValHisGlnAs 17
:::|||||
17 GTGCCCAAGAAAGAGAGAAAGTCTCCAACTCTGACTGCTGCACCAAAA 66
17 nLeuPrAlaLeuProValAspAlaThrSerAspGluValArgLysAsnL 34
|||||
67 CCTCCCTGCTCCCTCTGATGCGACCTCTGATGAGTCAGCAAGAAC 116
|||||
34 eumetAspMetPheArgAspArgGlnAlaPheSerGlnHisThrTrpLys 50
|||||
117 TGATGACATGTTTCAGGAGACAGCGCTTCTCTGAACACACCTGGAAG 166
|||||
51 MetLeuLeuSerValCysArgSerTrpAlaAlaTrpCysLysLeuSns 67

```

```

|||||
167 ATGCTCTGCTCTGTGTCAGATCTCGGCTGCTGCTGCAAGCTGAACAA 216
|||||
67 nArgLysTrpPheProAlaGluProGluAspValArgAspTrpLeuLeuT 84
|||||
217 CAGGAATGTTCTCCCTGCTGAACCTGAGATGATGAGGAGTAACTCTCTGT 266
|||||
84 yLeuGlnAlaArgGlyLeuAlaValLysThrLleGlnGlnHisLeuGly 100
|||||
267 ACCGTCAAGCCAGAGCGCTGCTGTGTGAAGACATCCAAACAGACCTGGGC 316
|||||
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSe 117
|||||
317 CAGTCAACAATGCTGCGACAGAGATCTGGCGCTGCGCTTCCTGTCGAC 366
|||||
117 TAsnAlaValSerLeuValMetArgArgLysArgLysGluAsnValAspA 134
|||||
367 CAATGCTGTGCTGCTGTGATGAGGAGAAATCAGAAAGAGAAATGTGATG 416
|||||
134 IagLysGluArgAlaLysGlnAlaLeuAlaPheGluArgTrpAspPheAsp 150
|||||
417 CTGGGAGAGAGAGCCAGACAGGCGCTGCTTGAACGCACTGACTTGTGAC 466
|||||
151 GlnValArgSerLeuMetGluAsnSerAspArgCysGlnAspLysArgAs 167
|||||
467 CAAGTCAGATCCCTGATGAGAACTGACAGATGCCAGAGCAATCCAGAA 516
|||||
167 nLeuAlaPheLeuGlyIleAlaTrpAsnThrLeuLeuArgLleAlaGluI 184
|||||
517 CTTGGCTTCTCTGCGCATCTGCTACAAACCCCTGCTGCGCATTCGCGAAA 566
|||||
184 IeAlaArgLleArgValLysAspLleSerArgThrAspGlyGlyArgMet 200
|||||
567 TTGCCAGATCAGAGTGAAGACATCTCCCGACCGATGGTGGAGATG 616
|||||
201 LeuIleHisLleGlyArgTrpThrLysThrLeuValSerThrAlaGlyVal 217
|||||
617 CTGATCCACATTTGCGAGACCAAGACCTGCTGTCCACAGTGGTGTGGA 666
|||||
217 uLysAlaLeuSerLeuGlyValThrLysLeuValGluArgTrpLleSerV 234
|||||
667 GAAGGCGCTGCTCCCGGGGGTTACCAACCTGCTGGAGAGATGATCTCTG 716
|||||
234 aLseGlyValAlaAspAspProAsnAsnTrpLeuPheCysArgValArg 250
|||||
717 TGTCTGTGTGCTGATGACCCCAACAACACTACCTGTTGCGCGGCTCAGA 766
|||||
251 LysAsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAl 267
|||||
767 AAGATGCTGTGCTGCCCTCTCTGCACTCCCAACTGTCACCCGCGGC 816
|||||
267 aLeuGlnGlyLlePheGlnAlaThrHisArgLeuLeuTrpGlyAlaLysA 284
|||||
817 CTTGAAGGAGATCTTTGAGGCCACCCACCGCTGATGTCATGATGTCAGAG 866
|||||
284 sPAspSerGlyGlnArgTrpLeuAlaTrpSerGlyHisSerAlaArgVal 300
|||||
867 ATGACTCTGGGCAAGATACCTGCGCTGCTGCGCACATCTCCAGAGTGTG 916
|||||
301 GlyAlaAlaArgAspMetAlaArgAlaGlyValSerLleProGluLleMe 317
|||||
917 GGTGCTCCAGGAGACATGAGCGAGGCTGTGTGTCATCTCCGAATATAT 966
|||||
317 tGlnAlaGlyGlyTrpThrAsnValAsnLleValMetAsnTrpLleAlaG 334
|||||
967 GCAGGCTGTGTGCTGAGCAATGTGAACATAGTGAAGAACTACATCAGAA 1016
|||||
334 snLeuAspSerGluTrnGlyAlaMetValArgLeuLeuGlnLyspLysp 350
|||||
1017 ACCTGAGACTCTGAGACTGGGCGCATGTGATGAGCTGCTCAGATGAGGAG 1066
|||||
seq_name: gb_pat:AX205077

```

```

seq documentation_block:
LOCUS AX205077 5261 bp DNA linear
DEFINITION Sequence 7 from Patent WO0155361.
ACCESSION AX205077
VERSION AX205077.1 GI:15394318
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 5261)
AUTHORS Hardy, S.F.
TITLE Recombinant aav packaging systems
JOURNAL Patent: WO 0155361-A 7 02-AUG-2001;
Chiron Corporation (US)
FEATURES
source
1..5261
location/qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
/note="recombinant DNA"
BASE COUNT 1309 a 1238 c 1329 g 1285 t
ORIGIN
alignment_scores:
Quality: 1792.00 Length: 349
Ratio: 5.135 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-662-128A-2 x AX205077 ..
Align seg 1/1 to: AX205077 from: 1 to: 5261
2 ProlyslsLysArgLysValSerAsnLeuThrValHisGlnAsnLe 18
1201 CCAAGAAGAAAGAGAAAGTTTCGATTTACTGACCGACCAAAATTT 1250
18 uProAlaLeuProValAspAlaThrSerAspGluValArgLysAsnLeuM 35
1251 GCCTGCATTACCGGTCGATGCACGAGATGAGAGATTTCGCAAGAACTGA 1300
35 eLaspMetPheArgAspArgGlnAlaPheSerGlnHisThrTryptLysMet 51
1301 TGGACATGTTCCAGGAGATCGCCAGCGCTTTCGTGACATACCTGGAATG 1350
52 LeuLeuSerValCysArgSerTrrAlaAlaTrrCysLysLeuAsnAsnAr 68
1351 CTTCTGTCGCGTTTCGCCGTCGTGGCGGCATGGCGCAATGGTAATACCG 1400
68 gLysTrrPheProAlaGluProGluAspValArgAspTyrLeuLeuTyrL 85
1401 GAAATGGTTCCCGCCAGACCTGAAGATGTTCCGATTAATCTTATATC 1450
85 euGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGlyGln 101
1451 TTTCAGGCGCGCGGCTTCGGCAGTAAAACTATCCACGAACTTTGGCGCAG 1500
102 LeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSerAs 118
1501 CTAAACATGCTTCATCGTCGTCGCGGCTCCACGACCAAGTACAGCAA 1550
118 nAlaValSerLeuValMetArgArgIleArgGlyGluAsnValAspAlaG 135
1551 TGCCTGTTTCCTGTTATGCGGGGATCCGAAAAAGAAAGCTGATGCGCG 1600
135 LysLysArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheAspGln 151
1601 GTGAACGTGCAAAACAGGCTCTACGCTTCGACAGCACTGATTTTCGACCA 1650
152 ValArgSerLeuMetLysAsnSerAspArgCysGlnAspIleArgAsnLe 168
1651 GTTCGTTCACTCATGGAAAAATACGATCCCTGCACAGATATACGTAATCT 1700

```

[illegible]

AUTHORS Bunting,M., Bernstein,K.E., Greer,J.M., Capecchi,M.R. and Thomas,K.R.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUL-1999) Human Genetics, University of Utah, 15 North 2030 East Room 5440, SLC, UT 84112-5331, USA
 FEATURES Location/Qualifiers
 Source 1.6652

/organism="Cloning vector pACN"
 /db_xref="taxon:100905"
 complement(682..715)
 /note="loxP element"
 polyA_signal complement(742..789)
 /note="HSV thymidine kinase polyA signal"
 gene complement(797..1600)
 /gene="neo"
 complement(797..1600)
 /gene="neo"
 /codon_start=1
 /product="neomycin phosphotransferase"
 /protein_id="AAD47089.1"
 /db_xref="GI:5690441"
 /translation="MGSATIEDGLHAGSPAAMVERLFGYDAAOQTGCSDAANVRLSA
 OGRPVLFYKTDLSGALNEDDEARLSWLTATGVCFAVLVDVTEAGRWLLGEVPG
 ODLSHLPAAKVSITMADAMRLTLPATCPEDHOKHRIEARTMEGLVDOD
 LDEHOGIAPALFRLKAKMPDGEDLVVTHGDCLPIMVENGSRFSGLDGRIGVA
 DRYODIATATRDIAELGSEWADRLVLYGIAAPDSORIAFYRLDEFF"
 complement(1617..2323)
 /note="RNA polymerase II large subunit promoter"
 complement(2339..2386)
 /note="HSV thymidine kinase polyA signal"
 complement(2533..3651)
 /gene="cre"
 complement(101n(2533..2736,2803..3651))
 /gene="cre"
 /codon_start=1
 /product="Cre recombinase with nuclear localization
 sequence"
 /protein_id="AAD47088.1"
 /db_xref="GI:5690440"
 /translation="MPKKRKVSNLITVHONLPALPYDATSDEVRKMLMDMPDROAF
 SEHTWKMLLVSCRSMAWCKLNNRKWPPEDEYRDLYLOARGLAVKTIQOHLGL
 NMLHRSGLPPSDSNVSLVMRIKRENDVAGRAKALAFERTDEPDVRSIMESD
 RCODIRNLAFGLIAYNTLRIATRIARVDISITDGRMLIHGRTKLTIVSTAGVEK
 ALSIGVTLERWISVSGVADDPNNYLFCVRKRGVAAPSAISQLSTRLEGITFEVTH
 RLIGAKDGGORYLANSGHARVGAARDMAKRGVSIPEITMAGWTNVIYNNYRN
 LDSEFGAVRLDEGDP"
 complement(2737..2802)
 /gene="cre"
 /note="SV40 t-antigen intron"
 complement(3652..4350)
 /note="testis specific angiotensin-converting enzyme
 promoter (tACE)"
 misc_feature complement(4358..4391)
 /note="loxP element"
 BASE COUNT 1593 a 1781 c 1700 g 1578 t
 ORIGIN

alignment_scores:
 Quality: 1776.00 Length: 372
 Ratio: 5.074 Gaps: 1
 Percent Similarity: 94.086 Percent Identity: 94.086

alignment_block:
 US-09-662-128a-2 x AF169416/rev ..

Align seg 1/1 to reverse of: AF169416 from: 1 to: 6652

1 MetProLysLysLysArgLysValSerasnLeuThrValHisGlnAs 17
 ||||||||||||||||||||||||||||||||||||||||||||
 3651 ATGCCCAAGAAAGAGAGAGAGATTACTGACCGACCAACAA 3602
 ||||||||||||||||||||||||||||||||||||||||||||
 17 nLeuProAlaLeuProValaAspAlaThrSerAspGluValArgLysAsnL 34

||||||||||||||||||||||||||||||||||||||||||
 3601 TTGGCTGCATTAACCGGTGATGCAACAGATGATGAGTTGCAAGAAC 3552
 ||||||||||||||||||||||||||||||||||||||||||||
 34 eUMetAspMetPheArgAspArgGlnAlaPheSerGluHisThrLys 50
 ||||||||||||||||||||||||||||||||||||||||||||
 3551 TGAATGACATGTTTCAGGATCGCCAGCGCTTTCTCGACATCACTCGAAA 3502
 ||||||||||||||||||||||||||||||||||||||||||||
 51 MetLeuSerValLysArgSerTrpAlaAlaTrpCysLysLeuAsnAs 67
 ||||||||||||||||||||||||||||||||||||||||||||
 3501 ATGCTTCGTGCTGCTTGGCGGTCGCGGCGGCGCATGTGCAAGTTGAATA 3452
 ||||||||||||||||||||||||||||||||||||||||||||
 67 nArgLysTrpPheProAlaGluProGluAspValArgAspTyrLeuLeu 84
 ||||||||||||||||||||||||||||||||||||||||||||
 3451 CCGGAATGGTTTCCCGCAACCTGAAGATGTTCCGAGATTATCTCTAT 3402
 ||||||||||||||||||||||||||||||||||||||||||||
 84 YLeuGlnAlaArgGlyLeuAlaValLysTrpTrpIleGlnHisLeuGly 100
 ||||||||||||||||||||||||||||||||||||||||||||
 3401 ATCTTCAGGCGCGCGCTTGGCAGATAAATATCCAGCAACTTTGGGC 3352
 ||||||||||||||||||||||||||||||||||||||||||||
 101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAsp 117
 ||||||||||||||||||||||||||||||||||||||||||||
 3351 CAGCTAAACATGCTTCATCGTCCGCGTCCGCGTGCACAGCAAGTACAG 3302
 ||||||||||||||||||||||||||||||||||||||||||||
 117 rAsnAlaValSerLeuValMetArgArgIleArgLysGluAsnValAsp 134
 ||||||||||||||||||||||||||||||||||||||||||||
 3301 CAATGCTGTTCACTGGTTATGCGCGGATCCGAAAGAAAGCTTGATG 3252
 ||||||||||||||||||||||||||||||||||||||||||||
 134 IArgLysAlaArgAlaLysGlnAlaLeuAlaPheGluArgTrpAspPheAsp 150
 ||||||||||||||||||||||||||||||||||||||||||||
 3251 CCGGTGAACGTCGCAAAACAGGCTCTACGTTCCAGACGACTGTTTCGAC 3202
 ||||||||||||||||||||||||||||||||||||||||||||
 151 GlnValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAs 167
 ||||||||||||||||||||||||||||||||||||||||||||
 3201 CAGGTGCTTCACTCATGAGAAATAGAGATCGCTCCAGAGATATACGTA 3152
 ||||||||||||||||||||||||||||||||||||||||||||
 167 nLeuAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuAlaGlyIle 184
 ||||||||||||||||||||||||||||||||||||||||||||
 3151 TCTGCACTTCTCGGGGATTCCTTATAACACCCGTTTACGATATGCCGAAA 3102
 ||||||||||||||||||||||||||||||||||||||||||||
 184 LeuAlaArgIleArgValLysAspIleSerArgThrAspLysGlyArgMet 200
 ||||||||||||||||||||||||||||||||||||||||||||
 3101 TTGCGAAGATCAGGCTTAAGATATCTCAGCTACGCTGCGGAGGAATG 3052
 ||||||||||||||||||||||||||||||||||||||||||||
 201 LeuIleHisIleGlyArgTrpThrLysThrLeuValSerThrAlaGlyVal 217
 ||||||||||||||||||||||||||||||||||||||||||||
 3051 TTATATCATATTTCGCAAGCAAGAAACGCTGTAGACCGCAGGTGAGA 3002
 ||||||||||||||||||||||||||||||||||||||||||||
 217 uLysAlaLeuSerLeuGlyValThrLysLeuValGluArgTrpIleSerV 234
 ||||||||||||||||||||||||||||||||||||||||||||
 3001 GAAGGCACTTAGCCCTGGGGTTAACTAACTGTCGACGCGATGATTTCCG 2952
 ||||||||||||||||||||||||||||||||||||||||||||
 234 aLserLysAlaAlaAspAspProAsnAsnTyrLeuPheCysArgValArg 250
 ||||||||||||||||||||||||||||||||||||||||||||
 2951 TCTCTGTGTAGCTGATGATCCGAATTAACCTACTGTTTCCGCGGTGAGA 2902
 ||||||||||||||||||||||||||||||||||||||||||||
 251 LysAsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAl 267
 ||||||||||||||||||||||||||||||||||||||||||||
 2901 AAAAATGCTGTCGCGCGCATCTCCACACCGCAGCTATCAACTACCGCGC 2852
 ||||||||||||||||||||||||||||||||||||||||||||
 267 aLeuGluGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAlaLys 283
 ||||||||||||||||||||||||||||||||||||||||||||
 2851 CCTGGAAGGATTTTGAAGCAACTCATCGATTGATTACGCGCGCTAAG 2802
 ||||||||||||||||||||||||||||||||||||||||||||
 283 283
 ||||||||||||||||||||||||||||||||||||||||||||
 2801 TAAATATAAATTTTAAGCTATAAGTGTTAAACTACTGATCTTAAT 2752
 ||||||||||||||||||||||||||||||||||||||||||||
 284 AspAspSerGlyGlnArgTyrLeuAlaTrpSerG 295
 ||||||||||||||||||||||||||||||||||||||||||||
 2751 GTTGTGATTTTAGAGATGACTGTGTCAGAGATACCTGCGCTGCTGG 2702
 ||||||||||||||||||||||||||||||||||||||||||||
 295 YHisSerAlaArgValGlyAlaAlaArgAspMetAlaArgAlaGlyValAs 312
 ||||||||||||||||||||||||||||||||||||||||||||

```

2701 ACACAGTCCCCGTGTCGAGCCGCGGAGATATGCCCCGCTGAGATT 2652
312 erllepogluilemetglnalaglytrphtasnvalasnileval 328
|||||
2651 CAAATCCGAGATCATGCAAGCTGTGGCTGGCCAAATGTAATATTGTC 2602
|||||
329 Metasntyrillearasnleuaspsergluthrghylametvalargle 345
|||||
2601 ATGAACATATTCGCTACCTGATAGTGAACAGGGCAATGTCGCCT 2552
345 uleugluaspglyasp 350
|||||
2551 GCTGGAAGATGGCGAT 2536

seq_name: gb_pat:AX150382

seq_documentation block:
LOCUS AX150382 1074 bp DNA linear PAT 25-JUN-2001
DEFINITION Sequence 1 from Patent EP118668.
ACCESSION AX150382
VERSION AX150382.1 GI:14571608
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
1 (bases 1 to 1074)
REFERENCE
AUTHORS Schwenk, F.
TITLE Transduction of recombinases for inducible gene targeting
JOURNAL Patent: EP 118668-A 1 25-JUL-2001;
ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES
Source
1..1074
/oranism="synthetic construct"
/db_xref="taxon:32630"
/note="DNA sequence coding for a fusion protein Tat-Cre"
CDS
1..1074
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC42784.1"
/db_xref="GI:14571609"
/translation="MGYGRKKRRQRRRNSNLLIVHONILPALPVDTSDVKKNLMD
FRDROAFSEHTMRLSVCRSMAMCKLNNRKFPAEPEDVDYLLVQAKLAVKI
QOHGOLNMLHRRSGLRPSDSNAVSLVMRRIKENVDAGERAKALAEKRTDQVR
SIMENSDRCODIRNLAFGLIAITLRIAEIARIKVDLSRFDGGRMLIHRTDFIV
STAGVEKALSLIGVTKLVERMISVGVADDPNNVILFCRVKNGVAAPSAITSQISTALE
GIEPATIRLIYGAKKDSDGORYLXMSGHSARVGAARMANAGVSIDEIMQSGMTVNI
VMNTRMLDSEGTAMRLLDDP"
BASE COUNT 268 a 254 c 299 g 253 t
ORIGIN

alignment_scores:
Quality: 1765.00 Length: 348
Ratio: 5.086 Gaps: 0
Percent Similarity: 99.713 Percent Identity: 98.563

alignment block:
US-09-662-128a-2 x AX150382 ..

Align seq 1/1 to: AX150382 from: 1 to: 1074

3 LysLysLysArgLysValSerAsnLeuLeuThrValHisGlnAsnLeu 19
|||||
28 CAACGCCGCGCGCATGTCATTAAGCTACCGCTACACCAAAATTTGCC 77
|||||
19 GAlaLeuProValAspAlaThrSerAspGluValArgLysAsnLeuMet 36
|||||
78 TGCATTTCCGGTCGATGCAACGAGATGAGGTCGCAAGAACCTGATGG 127
|||||
36 spMetPheArgAspArgGlnAlaPheSerGluHisThrTrpLysMetLeu 52
|||||
128 ACATGTTCAAGGATCGCAGCGCTTTTCTGAGCATACCTGAAATGCTT 177

```

```

53 LeuSerValCysArgSerTrpAlaIleTrpCysLysLeuAsnArgLys 69
|||||
178 CTGCTGGTTTGGCGGTGCTGGCGGATGTCAGTGAATGAACCGGAA 227
|||||
69 sTrpPheProAlaGluProGluAspValArgAspTrpLeuLeuTrpLeu 86
|||||
228 ATGTTTCCCGCAGAACCTGGAAGATGTTGCGATATATCTTATATCTTC 277
|||||
86 lnaIaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGlyLeu 102
|||||
278 AGCGCGCGGTCTGCAGTAAACATATCCAGCAACATTGGGCCAGCTA 327
|||||
103 AsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSerAsn 119
|||||
328 AACATGCTTATCTGCTCGCTCGGGCTGCCACGACCAAGTACGCAATGC 377
|||||
119 aValSerLeuValMetArgArgIleArgLysGluAsnValAspAlaGly 136
|||||
378 TGTTCACCTGGTTATCGCGCGGATCCGMAAAGAAAGTTGATGCCGGTG 427
|||||
136 lArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheAspGlnVal 152
|||||
428 AACGTCAAAACAGGCTCTAGCTTGCAGACGACCTGATTTGCACAGTT 477
|||||
153 ArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAsnLeu 169
|||||
478 CGTTCACCTCATGGAATAATAGCATCGCTCCAGAGATATACGTATCTG 527
|||||
169 aPheLeuGlyIleAlaLysAsnThrLeuLeuArgIleAlaGluIleAla 186
|||||
528 ATTCTGTGGGATTCCTTATACACCCCTGTTACGTAACGGAATGGCA 577
|||||
186 rGlyIleArgValLysAspIleSerArgThrAspGlyArgMetLeu 202
|||||
578 GGATCAGGGTTAAAGATATCTCACGTACTGACGCTGGGAAATGTAAATC 627
|||||
203 HisIleGlyArgThrLysThrLeuValSerThrAlaGlyAlaGluVal 219
|||||
628 CATATTGGCAGACGAAACGCTGTTAGCCACGCAAGTTAGAGAAAGC 677
|||||
219 aLeuSerLeuGlyValThrLysLeuValGluArgTrpIleSerValSer 236
|||||
678 ACTTAGCCTGGGGGTAACTAACTGCTGACGATGATTTCCGCTCTG 727
|||||
236 lValAlaAspAspProAsnAsnTrpLeuPheCysArgValArgLysAsn 252
|||||
728 GTGTAGCTGATGATCCGATTAACCTGCTTTTGCAGGGTCAGAAAAAT 777
|||||
253 GlValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAlaLeu 269
|||||
778 GGTGTTGCCCGCCCATCTGCCACGACGACATCACTCGCGCTGGA 827
|||||
269 uGlyIlePheGluAlaThrHisArgLeuIleTyrGlyValAlaLysAsp 286
|||||
828 AGGATTTTGAAGCAACTCATGATGATTTAGCGCCCTAAGGATGACT 877
|||||
286 eArgLysGlnArgTrpLeuAlaThrSerGlyHisSerAlaArgValGly 302
|||||
878 CTGCTCAGAGATACCTGGCTGCTGCGACACAGTCCCGCTGSGAGCC 927
|||||
303 AlaArgAspMetAlaArgAlaGlyValSerIleProGluIleMetGln 319
|||||
928 GCCGAGATATAGCCCGCGCTGAGTTCAATACCGGAGATCATGCAAC 977
|||||
319 aGlyGlyTrpThrAsnValAsnIleValMetAsnTrpIleArgAsnLeu 336
|||||
978 TGGTGGCTGCACCAATGTAATATTGTCATGAACATATTCCTTAACCT 1027
|||||
336 spSerGluThrGlyAlaMetValArgLeuLeuGluAspGlyAsp 350
|||||
1028 ATATGTAAACAGGGGCAATGGTGGCTGCTGGAAGATGGCGAT 1071

```


Integrated dna sequence
Patent: WO 0179512-A 23 25-OCT-2001;
Iconix Pharmaceuticals Inc. (US)

FEATURES
source
1. 4491
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector"

BASE COUNT 1160 a 1065 c 1119 g 1147 t
ORIGIN

alignment_scores:
Quality: 1765.00 Length: 348
Ratio: 5.086 Gaps: 0
Percent Similarity: 99.713 Percent Identity: 98.563

alignment_block:
US-09-662-128a-2 x AX347590 ..

Align seg 1/1 to: AX347590 from: 1 to: 4491

```

3 LysLysLysArgLysValSerAsnLeuThrValHisGlnAsnLeuP 19
:::.....:|||||:|||||:|||||:|||||:|||||:|||||:
169 GGTGAGCGCGTGGCTCCAAATTACTGACGCGTACACCAAAATTGGCC 218
19 GAlaleuProValAspAlaThrSerAspGluValArgLysAsnLeuMet 36
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 TGCATTACCGGTCGATGCACGAGTGAGAGGTTCCGCAAGAACCTGATGG 268
36 sPmetPheArgAspArgGlnAlaPheSerGluHisThrTrpLysMetLeu 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 ACATGTTCCAGGATCCGAGCGGCTTTCTGACATCACTGGAAAAATGCTT 318
53 LeuSerValCysArgSerTrpAlaAlaTrpCysLysLeuAsnAsnArgLys 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 CTGTCCGTTGGCGGTGCTGGGCGGCGCATGTGCAGTTGAATTAACCGGAA 368
69 sTrpPheProAlaGluProGluAspValArgAspTrpLeuLeuTrpLeuG 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 ATGGTTTCCCGCAGACCTGAAAGATGTTCCGATTAATCTCTATATATCTTC 418
86 lnaAlaArgLysLeuAlaValLysTrpIleGlnHisLeuGlnLeu 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 AGCGCGCGGTCTGGCAGTAAACATATCCACACATTTGGGCGCAGCTA 468
103 AsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSerAsnAl 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 AACATGCTTCATCTGTCGGTCCGGGTCGCCAGCAAGTACAGCAATATGC 518
119 aValSerLeuValMetCArgArgLysGlyGlnAsnValAspAlaGlyG 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
519 TGTTCCTCTGTTATGCGGCGGATCCGAAAAAGAAACGTGATGCCGGTG 568
136 lnaArgAlaLysGlnAlaLeuAlaPheGlnArgThrAspPheAspGlnVal 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
569 AACGTCAAAACAGCGCTAGCTTCGAACGCACTGATTTTCACCAAGTGT 618
153 ArgSerLeuMetGlnAsnSerAspArgCysGlnAspIleArgAsnLeuAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
619 CGTTCACCTACATGAAATAGCGATCGCTGCCAGATATACGTAATGTGGC 668
169 aPheLeuGlyIleAlaLysAsnThrLeuLeuArgLysIleAlaGlnIleAla 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
669 ATTTCCTGGGATTCCTATAACACCGCTAGCTATAGCGAAATGCGCA 718
186 rGllleArgValLysAspIleSerArgThrAspGlyGlyArgMetLeuIle 202
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
719 GGATCAGGGGTTAAAGATATCTCAGCTACTGACGCTGGAGAAATGTTATATC 768
203 HisIleGlyArgThrLysThrLeuValSerThrAlaGlyValGluLysAl 219
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
769 CATATGGCAGACGAAACCGCTGTGTAGACCCGCAAGGTGTAGAGAAAGGC 818

```

```

219 aLeuSerLeuGlyValThrLysLeuValGluArgTrpIleSerValSerG 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
819 ACTTAGCGCTGGGGTAACTAACTAGTCCAGCAGATGAGTATTCGCTCTG 868
236 lYValAlaAspAspProAsnAsnTrpLeuPheCysArgValArgLysAsn 252
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
869 GTGAGCTGATGATCCGAATAACTACTGTTTGGCGGGTCAAGAAAAAT 918
253 GlYValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAlaLeuG 269
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
919 GGTGTCCCGCGCATCTGCCACACGACCTTTCACACTGGCGCCCTGGA 968
269 uGlyIlePheGluAlaThrHisArgLeuIleTrpGlyAlaLysAspAsp 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
969 AGGATTTTGAAGCAGCATCATGATTTACGGCGCTAGAGATGACT 1018
286 eRgLYGlnArgTrpLeuAlaTrpSerGlyHisSerAlaArgValGlyAla 302
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1019 CTGGTCAGAGATACCTGCGCTGTGTCGACAGTGCCTGTGTCGAGCC 1068
303 AlaArgAspMetAlaArgAlaGlyValSerIleProGluIleMetGlnAl 319
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1069 GCGCAGATATGCGCCCGCGCTGGAGTTTCATATCCGAGATCATGCAGAC 1118
319 aGlyGlyTrpThrAsnValAsnIleValMetAsnTrpIleArgAsnLeu 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1119 TGTGCTGCGACCAATGAAATATTGTCATGAACTATATCCGTAACCTGG 1168
336 sPserGluThrGlyAlaMetValArgLeuLeuGluAspGlyAsp 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1169 ATAGTAAACAGGGCAATGTGTGCGCTGTGCGAAGATGGCGAT 1212

```

seq_name: gb_pat:AX191652

seq_documentation_block:

LOCUS AX191652 1125 bp DNA linear PAT 15-AUG-2001

DEFINITION Sequence 18 from Patent WO0149832.

ACCESSION AX191652

VERSION AX191652.1 GI:15209840

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

ARTEMIS Pharmaceuticals GmbH (DE)

location/Qualifiers

1. 1125

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="DNA sequence coding for a fusion protein

TATcrestleptag"

1. 1122

/note="unnamed protein product"

/codon_start=1

/transl_table=1

/db_xref="GI:15209841"

/translation="MGYRRKRRRRRRMSNLTTHQNPALPYDATSPVRRNLMD

FRDQAFSEHTWKMILSVCRSWAMWKLNNRWKPAPEPVEDRYLILDAKGLTYKTI

QOHGOLNMLHRRSGLPSPDSNAVSLVMRIKREVDAGERKQALAEERDFDVR

SLMNSDRCODIRNLATGLAYNTLIRLIRYADISRTDGGRLIHIGRTKLV

STAGEKRLSIGYKLYERTSVSGVADDDNNYLFGRKNGVAASATSOUSTRALE

GIFATEHRLILIGANDSGQRIANSGHSAVRCARARAGVSIPEIMAGGWTNVI

VNNYIRINDSETGAMVRLLEDGDIISIRGSAAMRHPQFG"

CDS

```

BASE COUNT 278 a 265 c 316 g 266 t
ORIGIN

```

alignment_scores:


```

1074 ATGCTTCCGACAGAACCTGAGAGTTCGGATTACTTCTATATCTTC 1025
      86 1nnlaarqgilyleuvalaVallysthrilleinglnhisleuaglylnleu 102
      1024 AGGCGCCGCGCTGACAGTAAAAACTATCCAGACATTTGGCCAGTAA 975
      103 AsnmetleuHisargargserglyleuProargProserAspserasn1 119
      974 AACATGCTTATCGTCGCTCCGCGCTCCACGACCAAGTACACCAATGC 925
      119 aValSerleuValMetargargilearglysluasValaspalagly 136
      924 TGTTTCACTGTTATGGCGCGATCCGAAAAAGAAAGTTGATGCCGGTG 875
      136 luatrgalalysglnaleuValaphegluarqthrAspPheaspglVal 152
      874 AACGTGCAAAACAGGCTCTACGCTTCAGACGACTGATTTTCGACAGGTT 825
      153 ArgSerleuMetGluasnSerAspArgCysGlnAspIleArgasnleuAl 169
      824 CGTTTACTCATGAAATAGCGATCGCTGCCAGGATATACGTATATCTGGC 775
      169 apheleuglyllealatyrsnThrleuValargIlealaglnlealaa 186
      774 ATTCTGGGGATTGCTTATACACCCGTGTACGTATAGCCGAATTCGCA 725
      186 rgIleargVallyAspIleSerargthrAspGlyIatArgmetleuIle 202
      724 GGATAGAGGTAAAGATATCTCAGCTACTGAGGTGGAGATATATATC 675
      203 HisIleArgThrlysthrleuValSerThrIleaglyValGluLysAl 219
      674 CATATTGGCAGACGAAACACCTGTTAGCACCCAGGTGTAGAGAAAGGC 625
      219 aleuSerleuGlyValThrlysluValGluArgTrpIleSerValSerg 236
      624 ACTATGCGCGGGGTAACTAACTGTCGACGAGATTCCTCGTCTCG 575
      236 lyValAlaAspAspProAsnAsnlyleuPheCysArgValArgLysAsn 252
      574 GTGTACTGATGATCGAATATCTACTCTGTTTGGCGGGTCAGAAAAAT 525
      253 GlyValAlaAlaProSerAlaThrSerGlnleuSerThrArgAlaLeuGl 269
      524 GGTGTTCCGCGCGCATCTGCCACACGACGATCACTCGCGGCTTGA 475
      269 uGlyIlePheGlnAlaThrHisArgleuIleTyrGlyAlaLysAspAsp 286
      474 AGGGATTTTGAAGCAATCATCATGATTGATTACGGCGCTAAGATGACT 425
      286 erGlyGlnArgTyrleuAlaTyrPserGlyHisSerAlaArgValGlyAla 302
      424 CTGTGACGAGATATCTGGCTGGCTGTGACACAGTCCGCTGCGAGACC 375
      303 AlaArgspMetAlaAlaArgAlaGlyValSerIleProGlnIleMetGlnAl 319
      374 GCCGAGATATGCGCCGCTGAGATTTCATATACCGGAGATATCAAGC 325
      319 aglyGlyTrpThrAsnValasnIleValMetasnTyrIleArgasnleuA 336
      324 TGGGGCTGGACCAATGTAAATTTCTCATGAACATATATCCGTAACCTGG 275
      336 spSerGluThrGlyAlaMetValArgleuGluAspGlyasp 350
      274 ATAGTGAACAGGGGCAATGTGCGCTGCGTAAGATGGCGAT 231
seq name: gb_pat:AX191647
seq documentation_block:
LOCUS AX191647 1623 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 13 from Patent WO0149832.
ACCESSION AX191647
VERSION AX191647.1 GI:15209835

```

```

KEYWORDS
SOURCE      synthetic construct.
ORGANISM    artificial sequence
REFERENCE   1 (bases 1 to 1623)
AUTHORS     Schwenk, F.
TITLE       Transduction of recombinases for inducible gene targeting
JOURNAL     Patent: WO 0149832-A 13 12-JUL-2001;
            ARTEMIS Pharmaceuticals GmbH (DE)
            Location/Qualifiers
FEATURES
   source
       1..1623
           /organism="synthetic construct"
           /db_xref="taxon:32630"
           /note="DNA sequence coding for a fusion protein
             deltaVP22core-StrepTag"
           1..1620
           /note="unnamed protein product"
           /codon_start=1
           /transl_table=11
           /protein_id="CAC51200.1"
           /db_xref="GI:15209836"
           /translation="MASMTGGQOQMGRODPSTAPTRSKTPAOGARKLHFSTAPRPPDAP
             KTPRVAGFNKRVFCAVGRLAAMHARMAVOLMDMSPRDEDELINELGITTIRVYC
             EKNILIORANELVNPDDVADATATGTRGSAARPTERRPARSARRRVEGTE
             LGSSTPVMNSNDIQHSGRMSNLTIVHQNLPALPYDATSDVEKRLNMDPRROAFS
             EHTKRWLSVCSKMAKSKINKRKRPKREPVDRDYLILYLOARGLAVKTIQOHLGQIN
             MLHRSGLEPRSDSNASVLVMRIRIKRENVNDEGRARAKALAEFRTPDOVRSIMNSDR
             CODIRNLAFLLGIAYNTLRLIAETIRIKVDLSRTDGGMLIHGRRTVLSTAGVEKA
             LSIQVTKLIVERMISVSGVADDPNNYLFRCVRKNVAAFSATQSULSTRALGIFBATIR
             LIYGADSDGSGYRSLAMSGHSAVGAARMARAGVSIPIEMQAGMTVNVIMNTIRNL
             DSRTGAMVRLDGDGIEGRGSAMRHPQFG"
BASE COUNT      368 a      440 c      471 g      344 t
ORIGIN
alignment_scores:
   Quality: 1759.00      Length: 344
   Ratio: 5.113          Gaps: 0
   Percent Similarity: 100.000      Percent Identity: 99.419
alignment_block:
US-09-662-128A-2 x AX191647 ..
Align seg 1/1 to: AX191647 from: 1 to: 1623
7 LysValSerAsnleuLeuThrValHisGlnAsnleuProAlaLeuProva 23
      ::::|||||
538 CGCATGTCCTCAATTTACTGACCGCTACACCAAAATTTGGCTGCATTAACCGGT 587
23 lAspAlaThrSerAspGlyValArgLysAsnleuMetAspMetPheArg 40
      :|||||
588 CGATGCAACGAGTGAAGGTTCCGACAGACCTGATGACATGTTCAAGGG 637
40 sPArgGlnAlaPheSerGluHisThrTrpLysMetleuLeuSerValCys 56
      :|||||
638 ATGCCAGCGGTTTCTGTGACATACCTGGAATAATCTTCTGCTGTTTGC 687
57 ArgSerTrpAlaAlaLeuTrpCysLysleuAsnAsnArgLysTrpPheProAl 73
      :|||||
688 CGGTCTGTGGCGGCGATGTGTCAGATTGAATTAACCGGAATGTGTTCCCGC 737
73 agluProGluAspValArgAspTyrleuLeuTyrleuGlnAlaArgGlyL 90
      :|||||
738 AGAACCTGAAGATGTTCCGAGTATATCTTCTATATCTTCAAGCGCGGCTG 787
90 euAlaVallysthrilleinglnhisleuGlyGlnleuAsnmetleuHis 106
      :|||||
788 TGGCAGTAAAAACATATCCAGCAATTTGGCCAGCTAAACATGCTTCTCAT 837
107 ArgArgserGlyleuProArgProserAspSerAsnAlaValSerleuVal 123
      :|||||
838 CGTCGCTCGCGGCTGCCACGACCAAGTACAGCAAGATGCTGTTCACTGCT 887

```

```

123 1metargargileargylsgluasnvalasplaglyluarqalalysg 140
|||||
888 TATGGGGCGATCCGAAAAGAACGTTGATGCCGGTGAACGTGCACAAAC 937
|||||
140 1na1aleualaepheliuarqthrargpheaapgluvalargserleumet 156
|||||
938 AGGCTCAGAGCTTCGAACGCACTGATTTTCAGCAGGTTGCTTCACTCATG 967
|||||
157 gluasnseraspargcygsluaspilleargasnleualapheluclyti 173
|||||
988 GAAATATGCGATCCGCTCCAGATATACGTAATCTGGCATTTCTGGGGAT 1037
|||||
173 ealatyasnthrleuenuargilealagluilealaargileargval 190
|||||
1038 TGCTTATATACACCTGTTAGCTATAGCCGAAATGGCCAGATCAGGGTTA 1087
|||||
190 ysapilserargthrargpjlgylyargmetleuilehislleglyarg 206
|||||
1088 AAGTATCTCAGCTACTGAGGGTGGAGATGTAATCCATATTTGGCAGA 1137
|||||
207 ThrlysthrleuvalserthrarglyvalglulysalaleuSerleugl 223
|||||
1138 ACAGAAACGGTGGTGTAGCACCGAGGTGTAGAGAACGCACTTACCTGGG 1187
|||||
223 yalThrlysthrleuvalgluarqtrpileservalserglyvalalaaspa 240
|||||
1188 GGTAACTAAAGTGTGAGCGATGATGATTCCTCTCTGTTAGCTGATG 1237
|||||
240 sPProasnaantyrleuphecyargvalarglysasnlyvalalaala 256
|||||
1238 ATCCGAAATACCTACCTGTTTCCGGGTCAGAAAATGGTGTGGCCGCG 1287
|||||
257 ProserAlaThrSerGlnLeuSerThrargalaleugllylephagl 273
|||||
1288 CCATCTGCACACGACGACATCAACTCGCCCTGGAGAGGATTTTGA 1337
|||||
273 uAlaThrHlsArgLeuiletyrGlyAlaLysAspSerGlylnargt 290
|||||
1338 ACCAATCATGATGATTACGGGCTCAAGATGACTGTGCTAGAGAT 1387
|||||
290 yfLeuAlaTrpSerGlyHlsSerAlaargvalglualaalaargspmet 306
|||||
1388 ACCTGGCTGGTGTGACACAGTCCCGTGTGGAGCCGCGGAGATG 1437
|||||
307 AlaargalaglyvalserlleProgluilemetGlnalaglytyrph 323
|||||
1438 GCCCGCGCTGAGTTCAATCCGAGATCATGCAAGCTGTGGCTGGAC 1487
|||||
323 rasnvalasnlevalmetasnTyrlleargasnleuaspserGlyuthrg 340
|||||
1488 CAATGTAATATTTGTCAATGAACTATATCCCTAACCTGATAGTGAACAG 1537
|||||
340 1yalametvalargleuenuaspjlgyasp 350
|||||
1538 GGCAATGAGTGGCGCTCTCGAAGATGGCGAT 1569
|||||
seq_name: gb_pat:AX150386
seq_documentation_block:
LOCUS AX150386 2004 bp DNA linear PAT 25-JUN-2001
DEFINITION Sequence 5 from Patent EP118668.
ACCESSION AX150386
VERSION AX150386.1 GI:14571612
KEYWORDS
SOURCE
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2004)
ARTIFICAL SEQUENCE.
AUTHORS Schwenk,F.
TITLE Transduction of recombinases for inducible gene targeting
JOURNAL Patent: EP 118668-A 5 25-JUL-2001;
ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES
Location/Qualifiers

```

```

source
1. 2004
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="DNA sequence coding for a fusion protein vp22-Cre"
1. 2004
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC42792.1"
/db_xref="GI:14571613"
/translation="MTRSRVSKSGPREVPDEYEDLYTPSSGMAIPDPSPTSRGA
LQTRRGGEVRFVQYDESDYALYGSSDDDEHPVPRPRRVSQAVLSGQPARAP
PPAGSGAGRPPTTAPRAPRTORVATKARAARAETTRKSAQPEASALPDAPAST
APTSKTPAOGCLARKLHSTRAPRPDPAPPTPRVAGEKRVCAAVGRLAAMAARMAV
SAASRPTRPAPARSASRPDRPEVETLSTSPVMNSADIQHSGRVSNLLTYNON
LPALPVDATSDVEKRLNMDRFDRQAQSEHTWMLSVCSRMAWCKLNNRMFPAPED
EDVADYLLYLOARGLAVKTIQOHLGOLNMLHRSGLPSPDSNAVSLVMRRIRKEND
AGERAKQALAEERTDFDQVRSLSMNSDRGODIRNLAFICAYVTLIRIETIRIKVD
ISRDGRMLIHIGRTKTLVSTAGVEKALISGYTLVERHISVGVADDPNNYLCRV
RKNGVAPSAATSOISTRALGIEFATHRLITGAKDSDGRTYLAWSGHSARVGAARDMA
RAGVSTPEIMQAGMTVMVMTIRNLDETGAMVRLLEDGD"
BASE COUNT 419 a 608 c 592 g 385 t
ORIGIN

alignment_scores:
Quality: 1759.00 Length: 344
Ratio: 5.113 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.419

alignment_block:
US-09-662-128a-2 x AX150386 ..
Align seg 1/1 to: AX150386 from: 1 to: 2004

7 LysValSerAsnLeuThrValHisGlnAsnLeuProAlaLeuProVa 23
|||||
970 CGCATGCTCCAAATTTCAGCTACACCAAAATTTGCTGCTATACCGGT 1019
|||||
23 1AspAlaThrSerAspGluValArgLysAsnLeuMetAspMetPheArg 40
|||||
1020 CGATGCAACGAGTATGAGTGTCCAGAAACCTGATGACATGTTCCAGG 1069
|||||
40 sPARGlnAlaPheSerGlnHisThrTrpLysMetLeuLeuSerValcys 56
|||||
1070 ATCCGCAAGCGCTTTCTGAGCATACCTGGAATAATGCTTGTGCTGGTTCG 1119
|||||
57 ArgSerTrpAlaAlaTrpCysLysLeuAsnAsnArgLysTrpPheProAl 73
|||||
1120 CGGTCTGTGGGGCGCATGTGCAAGTTGAATACCGGAAATGGTTTCCCGC 1169
|||||
73 aGluProGluAspValArgAspTyrlleuLeuTyrlleuGlnAlaArgGly 90
|||||
1170 AGAACCTGAAGATGTTCCGATATATCTTCAATATCTTACGCGCGCGGTC 1219
|||||
90 eua1aValLysThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHis 106
|||||
1220 TGGCAGTAAAACTATCCACACATTTGGGCCAGCTTAAACATGCTTCAT 1269
|||||
107 ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuVa 123
|||||
1270 CGTGGTCCGGGCTGCCACGACCAAGTGAACCAATGCTGTTCACTGGT 1319
|||||
123 1metargargileargylsgluasnvalasplaglyluarqalalysg 140
|||||
1320 TATGGCGCGGATCCGAAAAGAACGTTGATGCCGGTGAACGTGCACAAAC 1369
|||||
140 1na1aleualaepheliuarqthrargpheaapgluvalargserleumet 156
|||||
1370 AGGCTCAGAGCTTCGAACGCACTGATTTTCAGCAGGTTGCTTCACTCATG 1419
|||||
157 gluasnseraspargcygsluaspilleargasnleualapheluclyti 173
|||||

```

```

|||||
1420 GAAATATAGCATCGCTGCCAGATATACGTAATCGGCAATTCGGGGAT 1469
173 ealatyrasnthrleuvalargllealagluillealargyleuval 190
|||||
1470 TGGTTATACACCCCTGTACGTATAGCCGAAATTCACGAGATCAGGGTTA 1519
190 ysaspliserargthrpspglylarymetleuilehisileglyarg 206
|||||
1520 AAGATATCTCAGCTACTGACGGTGGAGATATTATCATATTGGCAGA 1569
207 ThrlysthrleuvalserthralaglylglulysalaleuSerleuGI 223
|||||
1570 ACGAAAACGGTGGTGTAGCACCGCAGGTGTAGAGAGGCACTTGCCGTGG 1619
223 yvalthrlyleuvalgluargtripleSerValserglyValAlaAspA 240
|||||
1620 GGTAACTAAACGTGTCGAGGATGATTCGCTCTGCTGTAGCTGATG 1669
240 spproAsnAsnTYrLeuPheCysArgValArglyAsnGlyValAlaAla 256
|||||
1670 ATCCGATTAACCTACCTGTTTGGCGGGTCAGAAAAAATGGTGTGCCGCG 1719
257 ProSerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyIlePheGI 273
|||||
1720 CCATCTGCCACACCGCAGCTATCACTCGCGCCCTGGAGAGGATTTTGA 1769
273 uAlaThrHisArgPheIleTYrGlyAlaLysAspAspSerGlyGlnArgT 290
|||||
1770 AGCAATCATGATGATTGATTACGGCGCTAAGATGACTGTGTGCAGAGAT 1819
290 yrluAlaIATPserGlyHisSerAlaArgValGlyAlaAlaArgAspMet 306
|||||
1820 ACCTGCGCTGCTGACACAGTCGCCGTGCGAGCGCGCGGATATG 1869
307 AlaArgAlaGlyValSerIleProGluIleMetGlnAlaGlyIleTYrPth 323
1870 GCCCGCGCTGGAGTTTCAATACCGAGATCATGCAAGCTGATGGCTGAC 1919
323 rAsnValAsnIleValMetAsnTYrIleArgAsnLeuAspSerGluPthrg 340
|||||
1920 CAATGTAATATTTGTCATGAACATAATCCGTAACCTGGATGTAAACAG 1969
340 lValAlaMetValArgLeuGluAspGlyAsp 350
1970 GGGCAATGTCGCTGCTGGAAGATGGCGAT 2001
seq_name: gb_pat: AX191639
seq_documentation_block:
LOCUS AX191639 2004 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 5 from Patent WO0149832.
ACCESSION AX191639
VERSION AX191639.1 GI:15209828
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 2004)
AUTHORS
Schweik, F.
TITLE
Transduction of recombinases for inducible gene targeting
JOURNAL
Patent: WO 0149832-A 5 12-JUL-2001;
ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES
location/Qualifiers
source
1..2004
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="DNA sequence coding for a fusion protein VP22-Cre"
1..2004
/note="unnamed protein product"
CDS
1..2004
/codon_start=1
/transl_table=1
/protein_id="CAC51197.1"

```

```

/db_xref="GI:15209829"
/translation="MTRSRSVSKSPGREVPDEYEDLYTPSSGNASPDSPDTSRGA
LQTRSNQREDEVRVOYDSDVALYGVSSSEDEHEVEVTRPRPVSGAVALSGPGARAP
PPAGSGGAGRTTTPARA PRIORVATATKAPAPAAETTRGRSAQPESAALPDAPAST
APTRSTKTPAGDLKRLHFEETAPRNPDAPEPTPRVAGFNKRVCAAGRLAAHAAVAAY
QIWDMSRPOTDEDLNELGTTITRTVCCKMLDRAHEIYNPDVVOVDNATATRGH
SAASRPTERRAPARSRRPRRVEGETELGSTSPVWMSADIQHSGGMSMLTYHON
LPALPYDATDEYKRLMDMFRDQAFSEHTWKMLLSYCRSAAMCKLNNKRPAPB
EDYRDLYLYQAGLAVKTQQHLQGLNMLHRSGLPRPSDNNAVSLVWRIRKENV
AGERAKOALAEFERTDPOYRSLMENSDRCDIRNLAFGLIAYNLIRLAELIRKVD
ISRTDGRMLIHGRKTLVSTAGVEKALISYTKLIVERMISVSGVADDPNNYLCRV
RKNGVAPSATQSULSTRALEGTIEATHRLITCAKDSGORYLAWSGHSARVCAADMA
RAGVSIPEIMQAGWTNVNIVNVIKNNLDSFTGAMVRLLEDG"
BASE COUNT 419 a 608 c 592 g 385 t
ORIGIN

```

```

alignment_scores:
  Quality: 1759.00      Length: 344
  Ratio: 5.113          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 99.419

```

alignment_block:

US-09-662-128a-2 x AX191639

Align seq 1/1 to: AX191639 from: 1 to: 2004

```

7 LysValSerAsnLeuLeuThrValHisGlnAsnLeuProAlaLeuProVa 23
970 CGCATGTCCATTTTACGACCGTACACCAAAATTTCCGCAATTCACGGT 1019
23 lAspAlaThrSerAspGluValArglyAsnLeuMetAspMetPheArg 40
1020 CGATGCAACGAGTGATGAGGTGCGCAAGACCTGATGACATCTTCAGGG 1069
40 sPARGlnAlaPheSerGlnHisThrTYrPlyMetLeuLeuSerValCys 56
1070 ATCCGACAGCGCTTTTCGACATACCTGAAATGCTGTGTCCTGCTTTC 1119
57 ArgSerTRPAlaAlaIATPcyslyleuAsnAsnArgtyrPheProAl 73
1120 CGGTGTCGGCGGCGATGTCGAAATGAATACCGAAATGTTTCCCGC 1169
73 aGluProGluAspValArgAspTYrLeuLeuTYrLeuGlnAlaArgGlyL 90
1170 AGAACCTGAAGATGTCGCGATATCTTCTATATCTTCAGCGCCGCGGTC 1219
90 euAlaValIlysthrIleGlnGlnHisleuGlyGlnLeuAsnMetLeuHis 106
1220 TGGCAGTAAAAACTATCCAGACATTTGGCCAGCTTAACATGCTTCAT 1269
107 ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerleuVa 123
1270 CGTGGTCCGGGCTGCCAGCAAGTGAAGCAATGCTGTTCACATGCT 1319
123 lMetArgArgIleArglyGluAsnValAspAlaGlyGluArgAlaLysG 140
1320 TATGCGCGCGATCCGAAAAAGAAACGTTGATGCGGTGAACGTCAGAAC 1369
140 lAlaIleuAlaPheGluArgThrAspPheAspGlnValArgSerleuMet 156
1370 AGGCTTACGCTTCGACACGCACTGATTCGACACAGTTCGTTCACTCATG 1419
157 GluAsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyI 173
1420 GAAATATAGCATCGCTGCCAGATATACGTAATCGGCAATTCGGGGAT 1469
173 eAlaTYrAsnThrleuLeuArgIlealagluillealargyleuval 190
1470 TGGTTATACACCCCTGTACGTATAGCCGAAATTCAGGATCAGGGTTA 1519
190 ysaspliserargthrpspglylarymetleuilehisileglyarg 206
|||||

```

```
1520 AAGATATCTCAGCTACTGACGCTGGGAGAAATGTTAAATCCATATATGGCAGA 1569
207 ThrIysThrIeuValSerThrAlaGlyValGluLysAlaIeuSerLeuGI 223
|||||
1570 ACGAAAACGCTGCTGTAGCACCAGGCTGTAGAAAGGCACTTAGCCTGGG 1619
223 yValThrLysIeuValGluArgTrpIleSerValSerGlyValAlaAspA 240
|||||
1620 GGTAACTAAACTGTCGAGCGATGATTCGCTCTGTGTAGCTGATG 1669
240 sPProAsnAsnTyrlIeuPheCysArgValArgLysAsnGlyValAlaAla 256
|||||
1670 ATCCGAATTAAGTACCTGTTTGGCGGGTCAGAAAAATGTGTGCCGCG 1719
257 ProSerAlaThrSerGlnIeuSerThrArgAlaIeuGluGlyIlePheGI 273
|||||
1720 CCATCTGCCACGACGACGCTATCAACTCGCGCCTGGAAAGGATTTTGA 1769
273 uAlaThrHisArgLeuIleTyrglyAlaLysAspAspSerGlyGlnArgT 290
|||||
1770 AGCACTCATCGATGATTACGGCGCTAAGGATGACTCTGGTCAGAGAT 1819
290 yrlIeuAlaTrpSerGlyHisSerAlaArgValGlyAlaAlaArgAspMet 306
|||||
1820 ACCTGGCTGCTGCTGACACACTGCCCGTGTGGAGCCGCGGAGATATG 1869
307 AlaArgAlaGlyValSerIleProGluIleMetGlnAlaGlyTrpTh 323
|||||
1870 GCCCGGCTGGAGTTCAATACCGAGATCATGCAAGCTGTGGCTGGAC 1919
323 rAsnValAsnIleValMetAsnTyrlIeArgAsnIeuAspSerGluThrg 340
|||||
1920 CAATGTAATATATTGTCATGAACTATATCCGTAACTGGATGTAACACAG 1969
340 lYAlaMetValArgIeuIeuGluAspGlyAsp 350
|||||
1970 GGGCAATGTGTGGCTCTCGAAGATGGCGAT 2001
```


OM of: US-09-662-128a-2 to: N_Geneseq_032802:* out_format : pfs

Date: May 10, 2002 4:15 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framst_pzn.model -DEV=xlp  
-Q/cgn2.1/USPTO_spool/US09662128/runat_08052002.135154_28269/app_query.fasta.1.413  
-DB=N_Geneseq_032802 -OFMT=fastap -SUFFIX=rrng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FEAPOP=6.000 -FEAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct  
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09662128 -ECGN1_1_595 -NCPu=6 -ICPU=3 -LONGLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

```
Query: US-09-662-128a-2  
Query length: 350  
Database: N_Geneseq_032802:*  
Database sequences: 1736436  
Database length: 858457221  
Search time (sec): 257.240000
```

score_list:

```
Sequence Strd Orig ZScore EScore Len ! Documentation  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH41175 + 1797.00 3456.81 3 4e-184 1  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09280 + 1797.00 3438.78 3 0e-183 4  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09269 + 1797.00 3438.78 3 0e-183 4  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH04928 + 1797.00 3438.78 3 1e-183 5  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH19900 + 1792.00 3438.43 3 6e-183 22  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH26331 + 1792.00 3429.18 1 2e-182 5  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002A.DAT:AAH09257 + 1765.00 3394.57 9 9e-181 1  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002A.DAT:AAH16712 + 1765.00 3378.63 7 6e-180 44  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09265 + 1761.00 3386.30 2 9e-180 1  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09263 + 1759.00 3370.31 2 2e-179 4  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09259 + 1759.00 3378.34 7 9e-180 1  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09259 + 1759.00 3376.00 1 1e-179 2  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09268 + 1759.00 3375.72 1 1e-179 2  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09266 + 1759.00 3367.02 3 4e-179 4  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH62541 + 1758.00 3380.03 6 4e-180 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH10215 + 1757.00 3379.51 6 4e-180 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH58065 + 1757.00 3373.69 1 4e-179 17  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH05702 + 1757.00 3372.12 1 8e-179 20  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH10217 + 1757.00 3370.37 2 2e-179 2  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997A.DAT:AAH05703 + 1757.00 3359.74 8 6e-179 6  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH81258 + 1756.00 3377.37 9 0e-180 1  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH10216 + 1754.00 3364.36 4 6e-179 2  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH10218 + 1754.00 3364.36 4 6e-179 2  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH10219 + 1754.00 3364.36 4 6e-179 2  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH62541 + 1741.00 3347.10 4 4e-178 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH62542 + 1739.00 3343.22 7 2e-177 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH62542 + 1737.00 3339.33 1 2e-177 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH62543 + 1729.00 3337.41 1 5e-177 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH62546 + 1729.00 3328.85 6 3e-177 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH62544 + 1721.00 3308.35 6 3e-176 11  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH19899 + 1658.00 3187.83 3 2e-169 10  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH09264 + 1713.00 3149.95 2 3e-111 5  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998A.DAT:AAH07268 + 671.00 1279.92 6 0e-63 69  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998A.DAT:AAH21209 + 120.00 125.95 114 89 16  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH53956 + 116.00 201.98 0 0067 8  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH81364 + 106.00 182.61 0 0799 8  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH84533 + 106.00 182.61 0 0799 8  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH52553 + 106.00 182.61 0 0799 8  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH53343 + 105.50 181.75 0 0892 8  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH56133 + 104.00 178.73 0 1314 8
```

```
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT:AAH58957 - 104.00 145.13 9.77  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999A.DAT:AAH33565 + 103.50 176.46 0.1757  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH13509 + 101.00 155.80 2.49  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH18790 + 101.00 155.07 2.73  
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH41223 + 100.50 105.51 1.6e+03
```

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH41175

seq_documentation_block:

ID AAH41175 standard; DNA: 1050 BP.

AC AAH41175;

DT 23-AUG-2001 (first entry)

DE Mammalian Cre recombinase gene.

KW Mammalian; Cre recombinase; organ transplantation; gene therapy;

KW animal disease model; ds.

OS Mammalia.

PN JP2001086989-A.

PD 03-APR-2001.

PF 17-SEP-1999; 99JP-0264364.

PR 17-SEP-1999; 99JP-0264364.

PA (OSAU) UNIV OSAKA.

DR WPI: 2001-360321/38.

PT P-PSDB; AAB98695.

Mammalian type Cre recombinase gene modified so as to have a high
expression efficiency in a mammal by selecting a codon of high
frequency for use in the mammal for expressing Cre recombinase -

Claim 2; Page 7-9; 12pp; Japanese.

The present sequence is a mammalian type Cre recombinase gene. This
sequence can be modified so as to have a high expression efficiency in a
mammal by selecting a codon of high frequency for use in the mammal for
expressing Cre recombinase. The gene can be used in organ
transplantation, gene therapy and creation of animal disease models.

Sequence 1050 BP; 204 A; 383 C; 339 G; 124 T; 0 other;

alignment_scores:

Quality:	1797.00	Length:	350
Ratio:	5.134	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-662-128a-2 x AAH41175 ..

Align seg 1/1 to: AAH41175 from: 1 to: 1050

```
1 MetProlylsLysArgLysValSerAsnLeuLeuThrValHisGlnAs 17  
|||||  
1 ATGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 50  
|||||  
17 nleuprolaileupprovalaapalathseraspcluvalargylasn 34  
|||||  
51 CTTGCCCGCCCTGCGCGGAGCGCCAGCAGGAGAGAGAGAGAGAG 100  
|||||  
34 eumetaspmetpheargasparglinalaphesercluhtsthrp 50  
|||||  
101 TGATGAGACATGCTTCGCGAGCGCGAGCGCTTCACAGGAGAGAG 150  
|||||  
51 MetLeuLeuSerValCysArgSerTrpAlaAlaTrpCysLeuAsnAs 67
```

```

|||||
151 ATGCTGCTAGCGGTGGCGCAGCTGGGCGCTGTGTCACCTGACCAACTACAA 200
67 nArglyStrpPheProAlaGluProGluAspValArgAspTyrLeuLeu 84
|||||
201 CCGCAAGTGTGTTCCCGCGAGCCCGAGAGACGTGGCGCATCTGCTGT 250
84 yrlLeuGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGly 100
|||||
251 ACCTGCAGGCCCGCGCGCTGGCGGTGAAGACCATCCAGACAGCAGCTGGGC 300
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSe 117
|||||
301 CAGCTAAACATGCTGCAGCCGCCGAGCGGCTGCCCGCCCGCCAGGAGACG 350
117 rAsnAlaValSerLeuValMetArgArgIleArgLysGluAsnValAspA 134
|||||
351 CAACGCCGTGAGCGCTGTGATGCGCGCATCCGCAAGAGAGACGTGAGAG 400
134 laGlyLysArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheasp 150
|||||
401 CCGCGCAGCGCGCCAAAGAGCGCTGGCTTGAGGCGACCGACACTTCGAC 450
151 GlnValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAs 167
|||||
451 CAGGTGCGCAGCGCTGATGGAGAAACAGCAGACCGCTGCCAGACATCCGCA 500
167 nLeuAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGlu 184
|||||
501 CCTGGCCTTCCTGGGCATCGCCCTACCAACCCCTGCGCATCGCGGAA 550
184 leAlaArgIleArgValLysAspIleSerArgThrAspGlyArgmet 200
551 TCGCCCGCATCCCGGTGAAGACATCAGCGCGCACCGACGCGCGCGCAG 600
201 LeuIleHisIleGlyArgThrLysThrLeuValSerThrAlaGlyValG 217
601 CTGATCCACATCGCGCGCAGCAAGACCCGTGTGAGCACCGCGCGCTGGA 650
217 ulysAlaLeuSerLeuGlyValThrLysLeuValGluArgTPPIlaSeV 234
651 GAAGGCCCTTGGCTGGCGGTGACCAAGCTGGTGAGCGCTGTGATACGG 700
234 alSerGlyValAlaAspAspProAsnAsnTyrLeuPheCysArgValArg 250
701 TGAGCGCGCGTGGCGAGCAGCCCAACAACACTGTTCTGCGCGCGCG 750
251 LysAsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAl 267
751 AAGAACGCGCTGGCGCGCCCGCAGCCAGCCAGCCAGCTGAGCACCGGGC 800
267 alLeuGluGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAlaLys 284
801 CCTGGAGGCAATCTTCAGAGCCACCCACCGCTGATCTACGGGCCAAG 850
284 spAspSerGlyLysArgTyrLeuAlaTrpSerGlyHisSerAlaArgVal 300
851 ACGACACGCGCGCAGCGCTACCTGAGCGCGCAGCGCGCGCGCTG 900
301 GlyAlaAlaArgAspMetAlaArgAlaGlyValSerIleProGluIleMe 317
901 GGGCGCGCGCGCAGCATGGCGCGCGCGGTGAGCATCCCGAGAGATCAT 950
317 tGlnAlaGlyIleTyrThrAsnValAsnIleValMetAsnTyrIleArg 334
951 GCGAGCGCGCGCGCTGAGCAACATGTAACATCGTATCACTACATCCGCA 1000
334 snLeuAspSerGluThrGlyAlaMetValArgLeuLeuGluAspGlyAsp 350
1001 ACCTGGACAGAGAGACCGCGCATGTGTGCGCTGTGAGAGACGGGAGAC 1050
seq_name: /STD1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD09280

```

```

seq_documentation_block:
ID      AAD09280 standard; DNA; 4847 BP.
XX
AC      AAD09280;
XX
DT      12-SEP-2001 (first entry)
XX
DE      pGK-cre-pA vector DNA.
XX
KW      DNA recombinase domain; protein transduction domain; PTD;
KW      gene alteration; fusion protein; Human immunodeficiency virus;
KW      HIV; pGK-cre-pA vector; ds.
XX
OS      Unidentified.
XX
PN      WO200149832-A2.
XX
PD      12-JUL-2001.
XX
PF      05-JAN-2001; 2001WO-EP00060.
XX
PR      07-JAN-2000; 2000BP-0100351.
PR      10-NOV-2000; 2000EP-0124595.
XX
PA      (ARTE-) ARTEMIS PHARM GMBH.
XX
PI      Schwenk F;
XX
DR      WPI; 2001-441873/47.
XX
PT      Using site-specific DNA recombinase domain/protein transduction domain
PT      fusion proteins for inducing target gene alterations in organisms or
PT      cell cultures -
XX
PS      Example; Page 67-68; 85pp; English.
XX
CC      The present invention relates to use of fusion proteins comprising
CC      a site-specific DNA recombinase domain e.g. Cre and a protein
CC      transduction domain (PTD) e.g. the Human immunodeficiency virus
CC      (HIV) derived TAT peptide, for preparing an agent for inducing
CC      target gene alterations in a living organism or cell culture. The
CC      present invention also provides a method for inducing gene
CC      alterations in living organisms using the fusion proteins of the
CC      invention. The present sequence is a pGK-cre-pA DNA.
XX
SQ      Sequence 4847 BP; 1139 A; 1238 C; 1283 G; 1184 T; 3 other;

alignment_scores:
      Quality: 1797.00      Length: 350
      Ratio: 5.134          Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-662-128A-2 x AAD09280 ..

Align seg 1/1 to: AAD09280 from: 1 to: 4847

1 MetProLysLysArgLysValSerAsnLeuLeuThrValHisGlnAs 17
|||||
2775 ATGCCCAAGAAAGAGAGAGGTGTCCAATTACTGACCGTACACCAAAA 2824
17 nLeuProAlaLeuProValAspAlaThrSerAspGluValArgLysAsnL 34
|||||
2825 TTTCCTGCATTACCGGTGATGCAACGAGTATGAGGTTCGCAAGAAC 2874
34 eumetaSpmetPheArgAspArgGlnAlaPheSerGluHisThrTyrLys 50
|||||
2875 TGATGACATGTTTCAGGAGTCCGACAGGCGTTTCTGACCAATACCTG 2924
51 MetLeuLeuSerValCysArgSerTyrPAlaIleTyrCysLysLeuAsnAs 67
|||||
2925 ATGCTTCTGTCCGTTTGGCGGTGCGGCGCATGTGCGCATGTGCAAGTGA 2974

```

```

67 narglystrpPheProAlaGluProGluAspValArgAspTyrLeuLeuT 84
   |||||
2975 CCGAATATGGTTCCCGCAGAACCTGAGATGTCGCGATTATCTTCTAT 3024
   |||||
84 yrLeuGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGly 100
   |||||
3025 ATCTTAGGCGCGGCTGTGGCAGTAAATACTATCCAGCAATTTGGGC 3074
   |||||
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspse 117
   |||||
3075 CAGCTAAATGCTTATGTCGTGGTCCGGCTGCCAGCAGCAAGTGACG 3124
   |||||
117 rAsnAlaValSerLeuValMetArgArgIleArgLysGluAsnValAspA 134
   |||||
3125 CAATGCTGTTCACTGCTTATGCGGGGATCCGAAAAGAAAAGCTTGATG 3174
   |||||
134 lAgIyGluArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheasp 150
   |||||
3175 CCGGTGAACGTGCAAAACAGGCTTAGCGTTCGAACGCACTGATTTGCG 3224
   |||||
151 GlnValArgSerLeuMetGluAsnSerAspArgGlyGlnAspIleArgas 167
   |||||
3225 CAGGTTCGTTCACTGATGAAAAATAGCCATCCCTCCAGGATATAGCTPA 3274
   |||||
167 nLeuAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGluI 184
   |||||
3275 TCTGGCATTTCTGGGATTTGCTTATACACCTGTTACGTATAGCCGAAA 3324
   |||||
184 lAlaArgIleArgValLysAspIleSerArgThrAspGlyIleArgmet 200
   |||||
3325 TTGCAGAGATCAGGTTAAAGTATCTCAGCTGACGACGAGTGAGAAATG 3374
   |||||
201 LeuIleHisIleGlyArgThrLysThrLeuValSerThrAlaGlyValG 217
   |||||
3375 TTAATCCATATTTGGCAGAACCAAAACGCTGTTAGCACCGAGGTGTA 3424
   |||||
217 uLysAlaLeuSerLeuGlyValThrLysLeuValGluArgTrpIleSerV 234
   |||||
3425 GAGGCACTTACCTCGGGGGTAACTGCTGAGCAGAGATGATTCGCG 3474
   |||||
234 aISerGlyValAlaAspAspProAsnAsnTyrLeuPheCysArgValArg 250
   |||||
3475 TCTGTGGTGTAGTGTGATGATCCGAATACTACTGTTTGGCGGGTCA 3524
   |||||
251 LysAsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgA 267
   |||||
3525 AAAATATGGTGTGGCGGCGCATCTGCCACAGCCAGCTATCAACTCCGC 3574
   |||||
267 aLeuGluGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAlaLysA 284
   |||||
3575 CCTGGAAGGGATTTTGAAGCACTCATCGATTGATTACGGCGCTAAGG 3624
   |||||
284 sPAspSerGlyGlnArgTyrLeuAlaTrpSerGlyHisSerAlaArgVal 300
   |||||
3625 ATGACTCTGTGCAGATACCTGCGCTGTGACACACTGCCCGCTGTC 3674
   |||||
301 GlYAlaAlaIArgAspMetAlaArgAlaGlyValSerIleProGluIleW 317
   |||||
3675 GGAACCCCGCGAGATATGGCCGCGGTGAGTTCAATACCGAGATCAT 3724
   |||||
317 tGlnAlaGlyGlyTyrThrAsnValAsnIleValMetAsnTyrIleArgA 334
   |||||
3725 GCAAGCGTGTGGCTGACCAATGTAAATATTGTCACTGAATATATCCGTA 3774
   |||||
334 sNeuAspSerGluThrGlyAlaMetValArgLeuLeuGluAspGlyAsp 350
   |||||
3775 ACCTGGATTACTGAACAGGGGCAATGTGCGCTGTGGAAGATGGCGAT 3824
   |||||
seq_name: /STD1/gcgdata/geneseq/geneseqn-embL/NA2001A.DAT:AAD09269
seq_documentation_block:
ID AAD09269 standard: DNA: 4960 BP.

```

```

XX
AC AAD09269;
XX
DT 12-SEP-2001 (first entry)
XX
DE pCMV-I-Cre-pA vector DNA.
XX
KW DNA recombinase domain; protein transduction domain; PTD;
KW gene alteration; fusion protein; Human immunodeficiency virus;
KW HIV; pCMV-I-Cre-pA vector; ds.
XX
OS Unidentified.
XX
PN W0200149832-A2.
XX
PD 12-JUL-2001.
XX
PE 05-JAN-2001; 2001MO-EP00060.
XX
PR 07-JAN-2000; 2000EP-0100351.
XX
PR 10-NOV-2000; 2000EP-0124595.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Schwenk F;
XX
DR WPI; 2001-441873/47.
XX
PT Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PT cell cultures -
XX
PS Example: Page 62-63; 85pp; English.
XX
CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived TAR peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a pCMV-I-Cre-pA vector DNA.
XX
SQ Sequence 4960 BP; 1225 A; 1213 C; 1296 G; 1226 T; 0 other;

```

alignment_scores:

	Quality: 1797.00	Length: 350
Percent Similarity: 100.000	Ratio: 5.134	Gaps: 0
	Percent identity: 100.000	

alignment_block:

```

US-09-662-128A-2 x AAD09269 ..
Align seg 1/1 to: AAD09269 from: 1 to: 4960

```

```

1 MetProLysLysLysArgLysValSerAsnLeuLeuThrValHisGlnAs 17
   |||||
977 ATGCCCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1026
   |||||
17 nLeuProAlaLeuProValAlaAspAlaThrSerAspGluValArgLysAsnL 34
   |||||
1027 TTTGGCTGTGATTACCGGTGATGATGCAACAGATGATGATGATGATGCAAGAAC 1076
   |||||
34 eUMetAspMetPheArgAspArgGlnAlaPheSerGluHisThrTrpLys 50
   |||||
1077 TATGAGACATGTTCCAGGATCCCGCAGGCTTTTCTGACCATACCTGGAAA 1126
   |||||
51 MetLeuLeuSerValCysArgSerTrpAlaAlaTrpCysLysLeuAsnAs 67
   |||||
1127 ATGCTTCTGTCCGTTGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1176
   |||||
67 nArgLysTrpPheProAlaGluProGluAspValArgAspTyrLeuLeuT 84

```

```

1177 CCGGAAAGGTTTCCCGCAGACCTGAGATGTTCCGATTATCTTCTAT 1226
      |||||||
84  YLEUGLINALARGLYLEUALVALYSTRHILEGLINHLISLEUGLY 100
      |||||||
1227 ATCTTCAGAGCCCGCGTCTGGCAGTAAATACTATCCAGCAACATTGGGC 1276
      |||||||
101 GINLEUASNMETLEUHLISARGISERGLYLEUPROARGPROSERASPSE 117
      |||||||
1277 CAGCTAAACATGCTTCATGTCGGTCCGCGTCCACGACCAAGTGACAG 1326
      |||||||
117  RASNLALVALSERLEUVALMETARGARGLIARGLYSLUASNLVALASPA 134
      |||||||
1327 CAATCTGTTTCACTGTTATCCGGCGCATCCGAAAAAAGAGCTGATG 1376
      |||||||
134  IAGLYLUARGALALYSGINALALEUALAPHEGLIARGTHASPHASP 150
      |||||||
1377 CCGGTGAACGTCCAAACAGCGCTAGCGTTGCAAGCAGCATGATTTCGAC 1426
      |||||||
151  GINVALARGSERLEUMETGLUASNSERASPARGYSGINASPLIARGAS 167
      |||||||
1427 CAGGTTGCTTCACTCATGAAATAGCGATGCTCCAGGATATACGTAA 1476
      |||||||
167  NLEUALPHLEUGLYLEALATYRASNTHREULEUARGYLEALAGLUI 184
      |||||||
1477 TCTGCGATTTCGGGGATTGCTTATACACCCTGTAGCTATAGCCGAAA 1526
      |||||||
184  LEALLARGILEARGVALLYSASPLIESERARGTHASPLIYARGMET 200
      |||||||
1527 TTGCCAGATCAGGTTAAAGATATCTCAGTACTAGCGTGGAGAAATG 1576
      |||||||
201  LEULIHLISILEGLYARTHRYSRTHREUVALSERTHRALGLIVALGL 217
      |||||||
1577 TTAATCCATATTTGGCGAAGCGTGGTTAGCACCGCAGGTGTA 1626
      |||||||
217  ULYSALALSERLEUGLYVALTHRLYSLEUVALGLUARGTRPILESERV 234
      |||||||
1627 GAAGGCACTTAGCTGGGGTAACTAACTGTCGAGCATGATGATTTCG 1676
      |||||||
234  ALSERGLYVALALASPARPROASNSRTHREUPHECYSRGVALARG 250
      |||||||
1677 TCTCTGCTGAGCTGATGATCCGAATACCTACTGTTTGGCGGCTACAA 1726
      |||||||
251  LYSASNGLYVALALALAPROSERALATHRSERGLINLEUSERTHARGL 267
      |||||||
1727 AAAAATGGTGTTCGCCGCATCTGCCACAGCCACTATCAACTGCGC 1776
      |||||||
267  ALDEUGLUGLYLEPHEGLUALATHRHISARGLEULIETRYRGYALALYSA 284
      |||||||
1777 CCTGGGAAGGATTTTGAAGCAACTCATCATGATTGATTTACGGCGTAAAG 1826
      |||||||
284  SPASPSERGLYGINARGTYRLEUALATRPSEGLYHISSEALAAARGVAL 300
      |||||||
1827 ATGACTCTGTCAGAGTACTGCTGCTGCTGACAGATGCGCTGTC 1876
      |||||||
301  GLYALALALARGASPMETALARGALGLYVALSERLEPROGLULEME 317
      |||||||
1877 GGAGCCCGCGAGATATAGCCCGCTGGAGTTTCAATACCGGAGATCAT 1926
      |||||||
317  TGINALAGLYLYTRPFRASNLVALSNILEVALMETASNTYRILEARGA 334
      |||||||
1927 GCAAGCTGGTGGCTGGACCAATGTAAATATTGTCAATGAACTAATATCCGTA 1976
      |||||||
334  SNLEUASPSERGLUTHRGYALAMETVALARGLEULEUGLUAASPLIYASP 350
      |||||||
1977 ACCTGGAATAGTGAACAGGGGCAATGTGCGCTGCTGGAAGATGGGAT 2026
      |||||||
seq_name: /STDS1/gcgdata/geneseq/geneseqn-emb1/NM2001A.DAT.AAD04928
seq_documentation_block:
ID  AAD04928 standard; DNA; 5365 BP.
XX
AC  AAD04928;

```

```

XX  17-JUL-2001 (first entry)
DE  Retroviral vector pBABE-pgkCre used as gene trap in embryonic stem cells.
XX
DE  Gene trapping construct; conditional mutation; unidirectional inversion;
XX  recombinase recognition sequence; RRS; disruption cassette;
XX  selection cassette; transgenic organism; retroviral vector;
XX  pBABE-pgkCre; cre recombinase; phosphoglycerate kinase; PGK promoter; ds.
OS  Chimeric - Moloney murine leukemia virus.
OS  Chimeric - Mus sp.
OS  Chimeric - Bacteriophage P1.
OS  Chimeric - Rhesus macaque polyoma virus.
XX
FH  Location/Qualifiers
FT  LTR
FT  8..480
FT  /tag= a
FT  /note= "5' partial MMLV U3, MMLV R and MMLV U5"
FT  /partial
FT  8..335
FT  /tag= b
FT  /note= "5' partial MMLV U3"
FT  /partial
FT  336..402
FT  /tag= c
FT  /rpl_type= DIRECT
FT  /note= "5' MMLV R direct repeat"
FT  403..480
FT  /tag= d
FT  /note= "5' MMLV U5"
FT  481..1374
FT  /tag= e
FT  /note= "MMLV primer binding site and extended
FT  packaging signal"
FT  1417..1921
FT  /tag= f
FT  /note= "Mouse phosphoglycerate kinase (PGK) promoter"
FT  1972..3024
FT  /tag= g
FT  /product= "Bacteriophage P1 Cre recombinase protein with
FT  a simian virus 40 large T-antigen N-terminal nuclear
FT  localisation signal"
FT  3088..3168
FT  /tag= h
FT  /note= "Promoter/enhancer deleted MMLV U3"
FT  3187..3332
FT  /tag= i
FT  /note= "3' MMLV R and MMLV U5"
FT  /partial
FT  3187..3253
FT  /tag= j
FT  /rpl_type= DIRECT
FT  /note= "3' MMLV R direct repeat"
FT  3254..3332
FT  /tag= k
FT  /note= "3' MMLV U5"
XX
XX  W0200129208-A1.
XX  26-APR-2001.
XX  PD
XX  16-OCT-2000; 2000WO-EP10162.
XX  PE
XX  16-OCT-1999; 99EP-0120592.
XX  PR  27-OCT-1999; 99US-0162016.
XX
XX  (ARTE-) ARTEMIS PHARM GMBH.
XX  (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX  Kuehn R, von Melchener H, Altschmied J;
XX  WPI; 2001-308486/32.
DR

```

XX New gene trapping construct capable of causing conditional mutations in
PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped

PS Example 1; Page 55-57; 78bp; English.

XX The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is retroviral vector PBAB-pgkcre which is
CC used as a gene trap in embryonic stem (ES) cells. The vector
CC PBAB-pgkcre is constructed by inserting Bacteriophage P1 cre recombinase
CC coding region with a simian virus 40 large T-antigen N-terminal nuclear
CC localisation signal and mouse phosphoglycerate kinase (PGK) promoter into
CC Moloney murine leukemia virus based vector PBABEpuro.

SQ Sequence 5365 BP; 1208 A; 1448 C; 1402 G; 1307 T; 0 other:

alignment_scores:

Quality: 1797.00 Length: 350
Ratio: 5.134 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-662-128A-2 x AAD04928

Align seg 1/1 to: AAD04928 from: 1 to: 5365

1 MetProLysLysArgLysValSerAsnLeuThrValHisGlnAs 17
1972 ATGCCCAAGAAAGAGAGAGAGTCCCAATTACTGACCCGACCAAAA 2021
17 nleuProLaleuProValaIaThrSerAspGluValArgLysAsnL 34
2022 TTGGCTGCATTACCGGTCGATGCAACGAGTGAAGTTCGCAAGAAC 2071
34 eumetAspmetPheArgAspArgGlnAlaPheSerGluHisThrTyrLys 50
2072 TGATGGACATGTTCAGGGATCGCCAGCGTTTCTGAGCATCTCTGAAA 2121
51 MetLeuSerValCysArgSerTrrPAlaIaIaIaIaIaIaIaIa 67
2122 ATGCTTCTGTCGTTTCCGCGTGTGGCGGCGCATGTGGCAAGTTGAATA 2171
67 nArgLysTrrPheProAlaGluProGluAspValArgAspTyrLeuLeuT 84
2172 CCGGAAATGGTTCCCGCAGACCTGAAGATGTTCGCGATTACTTCTAT 2221
84 YrLeuGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGly 100
2222 ATCTTCAAGGGCGGCGTGTGGCAGTAAACATATCCAGCAACATTTGGCG 2271
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspse 117
2272 CAGCTAAACATGCTTCATCGTGTCCGGCTGCGACGACCAAGTGCACG 2321
117 rAsnAlaValSerLeuValMetArgArgIleArgLysGluAsnValAspA 134
2322 CAATGCTGTTTCACTGTATATGCGGCGGATCCGAAAGAAAGAAACGTTGATG 2371

134 IagLyuArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheAsp 150
2372 CCGGTGAACGTGCAAAACAGGCTCTACCGTTCACAGCGACATGTTTCGAC 2421

151 GlnValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAs 167
2422 CAGGTGCTTCACATCATGAAATACGATCGCTGCCAGAGATATACGTAA 2471

167 nLeuAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGluI 184
2472 TCTGGCAATTTCTGGGATGCTTATTAACACCTGTACTGATATACCGAAA 2521

184 leuAlaArgIleArgValLysAspIleSerArgThrAspGlyLysArgmet 200
2522 TTGGCAGAGATCAGGTTAAAGATATCTACGTACTGACGTGGGGAATG 2571

201 LeuIleHisIleGlyArgThrLysThrLeuValSerThrAlaGlyValG 217
2572 TTTATCCATATTGGCAGAAACGAAACGGTGTAGCACCGCAGGTGTGA 2621

217 uLysAlaLeuSerLeuGlyValThrLysLeuValGluArgTrrPheSer 234
2622 GAAGGCACCTTAGCTGGGGGTAACTAACTGATGTCAGAGATGATTTCCG 2671

234 aISerGlyValAlaAspAspProAsnAsnTyrLeuPheCysArgValArg 250
2672 TCTCTGTGTACTGTATGATCGAATTAACCTGTTTGGCGGGTCA 2721

251 LysAsnGlyValAlaIaIaIaProSerAlaIaIaIaIaIaIaIaIa 267
2722 AAAAATGCTGTGCGCGCCGATCTCCACACGACGATATCACTCGCGC 2771

267 aLeuGluGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAlaLysA 284
2772 CCTGAGAGGATTTTGAAGCAACTCATCGATGATTTAGCGCGCTAAG 2821

284 sPAspSerGlyGlnArgTrrPheLeuAlaTrrPheSerGlyHisSerAlaArgVal 300
2822 ATGACTCTGTGTCAGAGATACCTGCTGTGTGGACACAGATGCGCTGTC 2871

301 GlyAlaAlaArgAspMetAlaArgAlaGlyValSerIleProGluIleMe 317
2872 GGAGCCGCGCAGATATGCGCGCGCTGAGTTTCATACCGAGATCAT 2921

317 tGlnAlaGlyIleTrrPheAsnValAsnIleValMetAsnTyrIleArgA 334
2922 GCAAGCTGTGTGCTGACCAATGTAAATATTGTCACTGAATATATCCGTA 2971

334 snLeuAspSerGluThrGlyAlaMetValArgLeuGluGluAspGlyAsp 350
2972 ACCTGAGATAGTGAACAGGGGCAATGTGCGCTGCGAAGATGGCGAT 3021

seq_name: /STDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX19900

seq_documentation_block:

ID AAX19900 standard; DNA; 2293 BP.

XX AAX19900:

DT 11-JUN-1999 (first entry)

DE Plasmid pOG304M containing mPl promoter and Cre coding sequences.

KW Mouse; mPl; promoter; site-specific recombination; Cre recombinase;

KM proteinase 1; fusion gene; transgenic mouse; ss.

OS Synthetic.

OS Mus musculus.

XX W09910488-A1.

XX PN

XX PD

XX 04-MAR-1999.


```
XX Adenovirus; delta-E1 Cre-LoxP Ad; adeno-associated virus; vector;
KW gene therapy; ds.
XX Chimeric - Mastadenovirus.
OS Chimeric - Human cytomegalovirus.
OS Chimeric - Rhesus macaque polyoma virus.
XX
FH Key Location/Qualifiers
FT promoter 551..1186
FT /tag= a
FT /note= "CMV promoter"
FT polyA_site 2251..2476
FT /tag= b
FT /note= "SV40 polyA site"
FT m1sc_recomb 2476..2520
FT /tag= c
FT /function= "loxP site"
XX
PN W0200155361-A2.
PD 02-AUG-2001.
XX
PF 26-JAN-2001; 2001WO-US02709.
XX
PR 26-JAN-2000; 2000US-0178536.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Hardy SF;
XX
DR WPI; 2001-483239/52.
XX
PT Producing recombinant adeno-associated virus (rAAV) vector, by stably
PT infecting eukaryotic host cell with rAAV vectors, helper adenovirus or
PT herpes virus, recombinant adenovirus or herpes vectors -
XX
PS Disclosure; Page 60-62; 63pp: English.
XX
CC- The present sequence is that of recombinant adenovirus delta-E1
CC- Cre-LoxP Ad. In which the E1 gene of the wild-type adenovirus is
CC- substituted by plasmid DNA comprising a cytomegalovirus (CMV)
CC- promoter inserted immediately downstream of the CMV promoter
CC- followed by a sequence from SV40 that specifies polyadenylation in
CC- mRNA, and a loxP site. The recombinant adenovirus is used as an
CC- induction system in methods of the invention. Thus, methods
CC- and compositions are provided for producing recombinant adeno
CC- associated virus (rAAV) vector particles by: (a) introducing into
CC- a host cell (1) AAV packaging plasmid p10xAAV (see AAH26326), (11)
CC- a recombinant viral vector encoding plasmid, and (111) a plasmid
CC- encoding herpesvirus, cytomegalovirus or adenoviral functions, or
CC- a herpesvirus, cytomegalovirus or adenovirus itself, to produce
CC- flox AAV particles and rAAV particles; and (b) introducing into a
CC- second host cell (1) the rAAV particles or (a), (11) a vector
CC- that directs expression of Cre, and (11) a vector which directs
CC- expression of herpes virus, cytomegalovirus or adenovirus helper
CC- functions, such that rAAV vector particles are produced. The
CC- vectors are useful for in vivo or in vitro gene therapy and also
CC- for in vitro recombinant protein production.
XX
SO Sequence 5261 BP; 1309 A; 1238 C; 1329 G; 1385 T; 0 other;

alignment_scores:
    Quality: 1792.00      Length: 349
    Ratio: 5.135          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment block:
US-09-662-128A-2 x AAH26331 ..
Align seg 1/1 to: AAH26331 from: 1 to: 5261
```

```
2 ProLysLysLysArGlyValSerAsnLeuThrValHisGlnAsnLe 18
|||||
1201 CCAAGAGAGAGAGAGAAAGTTTGAATTTACTGACCGTACACCAAAATTT 1250
18 UPtRoAlaLeuProValAspAlaThrSerAspGluValArgLysAsnLeu 35
|||||
1251 GCCTGCATTACCGTCGATGCCAAGAGTATGATGAGTTCGCAAGAACCTGA 1300
35 eAspMetPheArgAspArgGlnAlaPheSerGlnHisThrLysMet 51
|||||
1301 TGGACATGTTCAAGGATCCCAAGCCTTTTCTGACCATACCTGGAATATG 1350
52 LeuLeuSerValCysArgSerTrpAlaIatPrCysLysLeuAsnAr 68
|||||
1351 CTTCTGTCGGTTGGCCGTCGTGGGGGCGATGTCGAAGTTGAATACCG 1400
68 GlyStrPheProAlaGluProGluAspValArgAspTyrLeuLeuTyrL 85
|||||
1401 GAATGCTTCCCGCGAAGCTGAAGATGTTCGCGATTATCTCTATATTC 1450
85 euGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGlyln 101
|||||
1451 TTCAGCGCGCGCTCGGACGTAAAACTATCCAGCAACATTGGGCCAG 1500
102 LeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSer 118
|||||
1501 CTAAACATGCTTCATCGTCGTCGCGCTCCAGACCAAGTCACACCAA 1550
118 nAlaValSerLeuValMetArgArgIleArgLysGluAsnValAspAla 135
|||||
1551 TCCTGTTTCACGTGTTATCGCGCGGATCCGAAAGAAACGTTGATCCG 1600
135 LyGlnArgAlaLysGlnAlaLeuAlaPheGluArgTrhAspPheAspGln 151
|||||
1601 GTAAACGTCGCAAAACAGGCTCTACGTTCCGAACCACTGATTTCGACAG 1650
152 ValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAsnLe 168
|||||
1651 GTTCGTTCACTCATGGAATAAGCGATCCGCGCAGGATATACGTAAATCT 1700
168 uAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGluIle 185
|||||
1701 GGCATTCTCGGGATTCTTATACACCCGTGTTACGATAGCCGAATTC 1750
185 LaArgIleArgValLysAspIleSerArgThrAspGlyArgMetLeu 201
|||||
1751 CCAGGATCAGGGTTAAAGATATCTCACGTACTGACGTTGGAGAAATGTTA 1800
202 IleHisIleGlyArgThrLysThrLeuValSerThrAlaGlyValGly 218
|||||
1801 ATCCATATTGGCAGAAACGAAACGCTGTTAGCACCCGAGGTGTAGAGAA 1850
218 sAlaLeuSerLeuGlyValThrLysLeuValGluArgTrpIleSerVal 235
|||||
1851 GGCCTTAAGCCTGGGGTAACTAACTGGTCGACGATGATTCGCTCT 1900
235 eArgIleValAlaAspProAsnAsnTyrLeuPheGlyArgValArgLys 251
|||||
1901 CTGGTGTAGCTGATGATCCGAATTACTACCTGTTTGGCGGTCAGAAAA 1950
252 AsnGlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAlaLe 268
|||||
1951 AATGGTGTGTCGCGCCCATCTGCCACACCGCATATCAACTCGGGCCT 2000
268 uGluGlyIlePheGluAlaThrHisArgLeuLeuTyrGlyAlaLysAsp 285
|||||
2001 GGAAGGATTTTGAAGCAACTCATCGATGTATTACCGCCTTAAGATG 2050
285 sPserGlyLnaArgTyrLeuAlaIatPrSerGlyHisSerAlaArgValGly 301
|||||
2051 ACTGTGTCAGAGATACCTGCGCTGTGAGACACAGTCCCGTGTGCGA 2100
302 AlaAlaArgAspMetAlaArgAlaGlyValSerIleProGluIleMetGln 318
```

```

|||||
2101 GCCGGCGAGATATGCGCCGCTGAGTTTCATATACCGAGATCATGCA 2150
318 nalaglyglttrhrasnvalasnllevalmetasnryrileargasnL 335
2151 AGCTGGTGGCTGACCAATGTAATTTGTCTAGAACTATATCCGTAAAC 2200
335 euaspserrglutrhrglyalamelvalargleuenglualspg1asp 350
2201 TGGATATGTGAACAGGGGCAATGCTGCGCTGTGSAAGATGCGCAT 2247
seq_name: /SID1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAD09257
seq_documentation_block:
ID AAD09257 standard; DNA; 1074 BP.
XX
AC AAD09257;
XX
DT 12-SEP-2001 (first entry)
XX
DE TAT-Cre fusion DNA.
XX
KW DNA recombinase domain; protein transduction domain; PTD;
KW gene alteration; TAT-Cre fusion protein; Human immunodeficiency virus;
KW HIV; ds.
XX
OS Chimeric - Human immunodeficiency virus.
OS Chimeric - Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1..1074
FT /tag= a
FT /product= "TAT-Cre fusion protein"
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001MO-EP00060.
XX
PR 07-JAN-2000; 2000EP-0100351.
PR 10-NOV-2000; 2000EP-0124595.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Schwenk F.
XX
DR WPI: 2001-441873/47.
DR P-PSDB; AAE05264.
XX
PT Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PT cell cultures -
XX
PS Claim 17; Page 26-27; 85pp; English.
XX
CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived TAT peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a DNA encoding TAT-Cre fusion
CC protein. The TAT sequence is from HIV.
XX
SQ Sequence 1074 BP; 268 A; 254 C; 299 G; 253 T; 0 other;

```

```

alignment_scores:
  Quality: 1765.00      Length: 348
  Ratio: 5.086          Gaps: 0
Percent Similarity: 99.713  Percent Identity: 98.563

```

```

alignment_block:
US-09-662-128A-2 x AAD09257
Align seq 1/1 to: AAD09257 from: 1 to: 1074
3 LysAlaLysArgLysValSerAsnLeuThrValHisGlnAsnLeuP 19
28 CACACGGCGCGCGGCGATGTCATATTACTAGACCGTACACCAAAATTTGCC 77
19 oAlaLeuProValAspAlaThrSerAspGluValArgLysAsnLeuMet 36
78 TGCATTAACCGCTGATGACCAACGAGTGATGAGTTTGCAGAAACCTGATGG 127
36 spMetPheArgAspArgGlnAlaPheSerGlnHisThrTrpLysMetLeu 52
128 ACATGTTCAAGGATCCGACAGCGCTTTCTGAGCATCTGCAAAATGCTT 177
53 LeuSerValCysArgSerTrpAlaAlaTrpCysLysLeuAsnArgLys 69
178 CTGTCGCTTGCCGCTGCGTGGCGGCATGTGCAAGTTGAATTAACCGGGA 227
69 strPheProAlaGluProGluAspValArgAspTyrLeuLeuTyrLeuG 86
228 ATGGTTTCCCGCAGAACCTCAAGTGTCCGCAATTCCTTATATCTTC 277
86 InAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGlnLeu 102
278 AGCGCGCGGCTGCGCAGTAAACATATCCAGCAACATTTGGCCGCTA 327
103 AsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSerAsn 119
328 AACATGCTTCATCGTCCGCTCCGCGCTCCACGACCAAGTATACGCAATGC 377
119 aValSerLeuValMetArgArgIleArgLysGluAsnValAspAlaGly 136
378 TGTTCACCTGCTATGCGCGCGATCCGAAAGAAAGTTGATGCGGAG 427
136 LuArgAlaLysGlnAlaLeuAlaPheGluArgTrpAspPheAspGlnVal 152
428 AACGTGCAAAACAGGCTCTTACGCTTCAGCCACGACACTGATTCGACAGGTT 477
153 ArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAsnLeu 169
478 GGTTCACCATGAGAAATATAGCATCGCTCCGACGATATACGTAATCGGC 527
169 aPheLeuGlyIleAlaIleAsnThrLeuValSerThrAlaGluIleAla 186
528 ATTCTGGGATGCTTATACACCCCTGTACGTAATACCAATTCGCA 577
186 rGluLeuArgValLysAspIleSerArgTrpAspGlyArgMetLeuIle 202
578 GGTACGAGGTAAAGATATCTCAGCTACTGACGGGAGCAATGTATATC 627
203 HisIleGlyArgThrLysThrLeuValSerThrAlaGlyValGluVal 219
628 CATATTGGCAGAACGAAACGCTGTGTGACACCGCAGGTAGAGAAAGGC 677
219 aLeuSerLeuGlyValThrLysLeuValGluArgTrpLysSerValSerG 236
678 ACTTACCTGGGGGTACTAACTGTGTGAGGAGATTCCTCGCTCG 727
236 LyValAlaAspAspProAsnAsnTyrLeuPheCysArgValArgLysAsn 252
728 GTGAGCTGATGATCCGAATTAACCTCTTTTGGCGGCTGAGAAATAT 777
253 GlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAlaLeuG 269
778 GGTGTGCGCGCCATCTGACACCAACGCTATCACTACGCGCCCTGGA 827
269 uGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAlaLysAspAsp 286
828 AGGATTTTGAAGCAACTCATGATTAATTAACGCGCTAAGGATGACT 877

```



```

678 ACTTACCTGGGGTAACTAACTGTCGAGCGATGGATTTCGCTCTG 727
236 1YVALAAlaAspProAsnAsnTyrLeuPheCysArgValAlaGlyAsn 232
    |||||
728 GTGAGTGTGATGCCAATRACCTGCTTTGCGGGTCAGAAAAAT 777
253 G1YVALAAlaProSerAlaThrSerGlnLeuSerThrArgAlaLeuG1 269
    |||||
778 GGTGTTCGCCGCCATCTGCCACAGCAGCTATCAACTGCGCCCTCGGA 827
269 UG1YLePheGluAlaThrHisArgLeuLeuTyrG1YAlaAlaGlyAsp 286
    |||||
828 AGGCAATTTTGAAGCACTCATGATTGATTACGGCGCTAAAGAGTACAT 877
286 ERG1GlnArgTyrLeuAlaThrSerG1YHisSerAlaArgValG1YAla 302
    |||||
878 CTGGTAGAGATACCTGGCCTGCTGCGACACAGTCCCGCTGCGAGCC 927
303 ALAArgAspMetAlaArgAlaG1YAlaSerIleProGluIleMetGlnAl 319
    |||||
928 GCCGAGATATGCCCCCGCTGGAGTTTCATACCGAGATCATGCAAGC 977
319 AG1YG1YTPThrAsnValAsnIleValMetAsnTyrIleArgAsnLeu 336
    |||||
978 TGGTGGCTGACCAATCTAAATTTCTCATGACTATATCCGTAAACCTGG 1027
336 SPSErGluThrG1YAlaMetValArgLeuLeuG1uAspG1YAsp 350
    |||||
1028 ATAGTGAAACAGGGGCAATGTGGCCTGCTGGAAGATGGCGAT 1071

seq_name: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAD09265
seq_documentation_block:
ID   AAD09265 standard; DNA; 4727 BP.
AC   AAD09265:
XX
XX
XX   AAD09265:
XX
XX   12-SEP-2001 (first entry)
XX
XX   pT7-TACS vector DNA related to the invention.
XX
XX   DNA recombinase domain; protein transduction domain; PTD;
XX   gene alteration; fusion protein; Human immunodeficiency virus;
XX   HIV; pT7-TACS vector; ds.
XX
XX   Unidentified.
XX
XX   MO200149832-A2.
XX
XX   12-JUL-2001.
XX
XX   05-JAN-2001; 2001MO-EP00060.
XX
XX   07-JAN-2000; 2000EP-0100351.
XX   10-NOV-2000; 2000EP-0124595.
XX
XX   (ARTE-) ARTEMIS PHARM GMBH.
XX
XX   Schwenk F;
XX
XX   WPI; 2001-441873/47.
XX
XX
XX   Using site-specific DNA recombinase domain/protein transduction domain
XX   fusion proteins for inducing target gene alterations in organisms or
XX   cell cultures -
XX
XX   Example 1; Page 49-51; 85pp; English.
XX
XX   The present invention relates to use of fusion proteins comprising
XX   a site-specific DNA recombinase domain e.g. Cre and a protein
XX   transduction domain (PTD) e.g. the Human immunodeficiency virus
XX   (HIV) derived TAR peptide, for preparing an agent for inducing
XX   target gene alterations in a living organism or cell culture. The

```

```

CC   present invention also provides a method for inducing gene
CC   alterations in living organisms using the fusion proteins of the
CC   invention. The present sequence is a pT7-TACS vector DNA related to
CC   the invention.
XX
XX   Sequence 4727 BP; 1172 A; 1232 C; 1177 G; 1146 T; 0 other;

alignment_scores:
    Quality: 1761.00      Length: 348
    Ratio: 5.075          Gaps: 0
    Percent Similarity: 99.713      Percent Identity: 98.276

alignment_block:
US-09-662-128A-2 x AAD09265/rev ..

Align seg 1/1 to reverse of: AAD09265 from: 1 to: 4727

3  LysLysLysArgGlyValSerAsnLeuThrValHisGlnAsnLeuPr 19
   ::::::::::: :::::::::::
1274 CAACGCCGCCGCCGATGTCGAATTTACTGACCCGTACACCAAAATTTGCC 1225

19  OAlaLeuProValAspAlaThrSerAspGluValArgGlyAsnLeuMet 36
   :::::::::::
1224 TGCAATTACGGGTCGATGCCAGAGTGAAGAGTTGCAAGAACCTGATGG 1175

36  SPMetPheArgAspArgGlnAlaPheSerGluHisThrTyrLeuMetLeu 52
   :::::::::::
1174 ACATGTTCCAGGATCGCCAGCGCGTTTCTGAGCATCTGGAAGAAATGCTT 1125

53  LeuSerValCysArgSerThrAlaIleTyrCysLysLeuAsnAsnArgLy 69
   :::::::::::
1124 CTGTCCGTTTCCGCTGCTGGCGGCAGTGCGCAAGTTGAATTAACCGGAA 1075

69  STRPheProAlaGluProGluAspValArgAspTyrLeuTyrLeuG 86
   :::::::::::
1074 ATGGTTTCCCGCAGAACCTGGAAGATGTTCCGAGTATCTCTATATCTTC 1025

86  IuAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGlyGlnLeu 102
   :::::::::::
1024 AGGGCGCGGCTGACAGTAAACCTATCCAGCAACATTTGGCGCCAGCTA 975

103  AsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSerAsnAl 119
   :::::::::::
974 AACATGCTTCATCGTGGTCCGGGCTCCACGACCAAGTACAGCAATGC 925

119  aValSerLeuValMetArgArgIleArgGlyGluAsnValAspAlaGlyG 136
   :::::::::::
924 TGTTCACCTGTTATGGCGGATCCGAAAGAAAGATTGATGCCGGTG 875

136  LuArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheAspGlnVal 152
   :::::::::::
874 AACGTGAAAACAGGCTTAAGCTTTCGACGACCTAATTTGACACGAGTT 825

153  ArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAsnLeuAl 169
   :::::::::::
824 CGTTCACTCATGGAATAAGCAGTCCTGCGACAGATATACGTATCTGCG 775

169  aPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGluIleAla 186
   :::::::::::
774 ATTTCTGGGATTTGCTTATTAACACCTGTTACGTATAGCGAAATTCGCA 725

186  rG1LeArgValLysAspIleSerArgThrAspG1YG1YrMetLeuLe 202
   :::::::::::
724 GGATCAGGGTTAAAGATATCTCACGTAAGAGCGGTGGGGAATGTATATC 675

203  HisIleG1YArgThrLysThrLeuValSerThrAlaGlyValGluLysAl 219
   :::::::::::
674 CATATTGGAGAAACGAAACGCTGTTAGCACCGAGGTGTAGAGAAAGCC 625

219  aLeuSerLeuGlyValThrLysLeuValGluArgTrpIleSerValSerG 236
   :::::::::::
624 ACTTAGCCTGGGGGTAACTAACTGTCGAGCGATGATTCGCTCTG 575

```

```

236 1YValAlaAspAspProAsnAsnTyrLeuPheCysArgValAlaArgLysAsn 252
      |||||||
574 GTGATAGCTGATGATCCGATATACACTGTTTGGCGGGTCAGAAAAAAT 525
253 GLYValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAlaLeuGI 269
      |||||||
524 GGTGTGGCGCGCATCTGCACACGACCATCAACTCGCCCGCTGGGA 475
269 uGIYLePheGlnAlaThrHisArgLeuIleTyrGlyAlaLysAspAsps 286
      |||||||
474 AGGGATTTTGAAGCACTCATCATGATTGATTAGCGGCGCTAAGGATGACT 425
286 eRGlyGlnArgTyrLeuAlaTrpSerGlyHisSerAlaArgValAla 302
      |||||||
424 CTGGTCAGAGATACCTGCGTGTGTGACACACTGCCGCGTGGAGCC 375
303 AlaArgAspMetAlaArgAlaGlyValSerIleProGlnIleMetGlnAl 319
      |||||||
374 GCGCAGATATGGCCCGCGCTGGAGTTTCATACCGAGATCATGCAAGC 325
319 aGIYGIYTrpThrAsnValAsnIleValMetAsnTyrIleArgAsnLeuA 336
      |||||||
324 TGTGTGGCTGGACCAATGTAATATGTGATGACATATATCCGTAACCTGG 275
336 sPSerGIUThrGlyAlaMetValaArgLeuLeuGluAspGlyAsp 350
      |||||||
274 ATAGTGAAACAGGGGCAATGTGTGCGCTGTGGAAGATGGCGAT 231
seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD09263
seq_documentation_block:
ID AAD09263 standard; DNA; 1623 BP.
XX
AC AAD09263;
XX
DT 12-SEP-2001 (first entry)
XX
DE Delta VP22Cre-Streptag fusion DNA.
XX
KM DNA recombinase domain; protein transduction domain; PTD;
KM gene alteration; delta VP22Cre-Streptag fusion protein;
KM Human immunodeficiency virus; HIV; Human spumaretrovirus; HSV; ds.
XX
OS Chimeric - Human spumaretrovirus.
OS Chimeric - unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1..1620
FT /*tag= a
FT /product= "Delta VP22Cre-Streptag fusion protein"
XX
PN WO200149832-A2.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-EP00060.
XX
PR 07-JAN-2000; 2000EP-0100351.
PR 10-NOV-2000; 2000EP-0124595.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Schwenk F;
XX
DR WPI; 2001-441873/47.
DR P-PSDB; AAE05270.
XX
XX Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PT cell cultures -
XX
PS Claim 17; Page 44-46; 85pp; English.

```

```

XX
CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the human immunodeficiency virus
CC (HIV) derived Tat peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a DNA encoding delta
CC VP22Cre-Streptag fusion protein. The VP22 sequence is from
CC Human spumaretrovirus (HSV).
CC
SQ
Sequence 1623 BP; 368 A; 440 C; 471 G; 344 T; 0 other;

alignment_scores:
      quality: 1759.00      length: 344
      Ratio: 5.113      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.419

alignment_block:
US-09-662-128A-2 x AAD09263 ..
Align seg 1/1 to: AAD09263 from: 1 to: 1623

7 1YValSerAsnLeuThrValHisGlnAsnLeuProAlaLeuProVa 23
      :::::::::::
538 CGCATGTCCAATTTACTGACCGTACCCAAATTTGGCTCGCATTTACCGGT 587
23 1AspAlaThrSerAspGluValaArgLysAsnLeuMetAspMetPheArg 40
      |||||||
588 CGATGCAACGAGTATGAGTTTCGACAGAACCTGATGACATGTTCAGGG 637
40 sPArgGlnAlaPheSerGlnHisThrTrpLysMetLeuLeuSerValCys 56
      |||||||
638 ATCGCCAGGCGTTTCTGACATACCTGGAATAATGCTTGTGCTTTCG 687
57 ArgSerTrpAlaAlaTrpCysLysLeuAsnAsnArgLysTrpPheProAl 73
      |||||||
688 CGGTCTGGCGCGCATGTGTCAAGTTGAATACCGGAATGTGTTCCCGC 737
73 aGIUProGluAspValaArgAspTyrLeuLeuTyrLeuGlnAlaArgGlyL 90
      |||||||
738 AGAACCTGAAGATGTTCCGCGATTATCTCTATATCTTACGCGCGGTC 787
90 eValAlaValLysThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHis 106
      |||||||
788 TGCGCAGTAAAAACTATCCAGCAACATTTGGCCACCTAAACATGCTTCAT 837
107 ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuVa 123
      |||||||
838 CGTGGTCCGCGCTGCCACGACCAAGTACAGCATGCTGTTTCACGTGT 887
123 1MetArgArgIleArgLysGluAsnValaAspAlaGlyLuarGAlaLysG 140
      |||||||
888 TATGCGCGCGATCCGAAGAAAAACGTTGATCCGCTGAACGTGCAAAAC 937
140 1AlaIleuAlaPheGlnuArgThrAspPheAspGlnValaArgSerLeuMet 156
      |||||||
938 AGCCTTAGCGTTTCGAACGACGATGATTCGACCAAGTTCGTTCACTATG 987
157 GluAsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyI 173
      |||||||
988 GAAATATGCGATCGTCCGAGATATAGCTAAATTCGCACTTCTGGGAT 1037
173 eAlaTyrAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArgValL 190
      |||||||
1038 TGCTTATACACACCTGTTTACGTATAGCCGAAATTCGACAGATCAGGTTA 1087
190 yAspIleSerArgThrAspGlyGlyArgMetLeuIleHisIleGlyArg 206
      |||||||
1088 AAGATATCTACGTACTGACGCGTGGAGAAATGTTATTCATATTTGCAACA 1137

```

```

207 ThrLysThrLeuValSerThrAlaGlyValGluValAlaLeuSerLeuG1 223
1138 ACGAAAACGGCTGGTTAGCACCGAGCTGTAGAGAAAGCCTTAGCCTGGG 1187
223 YValThrLysLeuValGluArgTrpLysSerValSerGlyValAlaAspA 240
1188 GGTAACTAAAGCTGTCGAGCGATGATGATTCGCTCTGCTGTAGCTGATG 1237
240 SPProAsnAsnTrpLeuPheCysArgValAlaArgLysAsnGlyValAla 256
1238 ATCCGATTAATACACTACCTGTTTCCGGGTAGAAAATAGTGTCCCGCG 1287
257 ProSerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyLlePheG1 273
1288 CCATCTGCCACACGACGCTATCACTCGCGCCCTGGAAGGATTTTGA 1337
273 uAlaThrHisArgLeuLleTyrGlyValAlaLysAspSerGlyGlnArgT 290
1338 AGCAACTCATGATGATTTACGGCGCTAAGGATGACTCTGTGTCAGAGAT 1387
290 YLLeuAlaTrpSerGlyHisSerAlaArgValGlyValAlaArgAspMet 306
1388 ACCTGCCCGCTGTGACACACTGCCGCTGCGAGCCGCCGCGAGATATG 1437
307 AlaArgAlaGlyValSerLleProGluLleMetGlnAlaGlyLleTrpTh 323
1438 GCCCGCGCTGGAGTTTCATACCGGAGATCATGCAAGCTGCTGCTGAC 1487
323 rAsnValAsnLleValMetAsnTrpLleArgAsnLeuAspSerGluThrG 340
1488 CAATGTAATATGTCATGAACATAATCCGTAACCTGATGATGAAACAG 1537
340 YAlaMetValArgLeuLeuGluAspGlyAsp 350
1538 GGGCAATGCTGCGCTGCTGGAAGATGGCGAT 1569

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAD09259
seq_documentation_block:
ID AAD09259 standard; DNA; 2004 BP.
XX
AC AAD09259;
XX
DT 12-SEP-2001 (first entry)
XX
DE VP22-Cre fusion DNA.
XX
KW DNA recombinase domain; protein transduction domain; PTD;
KW gene alteration; VP22-Cre fusion protein; Human immunodeficiency virus;
KW HIV; Human spumaretrovirus; HSV; ds.
XX
OS Chimeric - Human spumaretrovirus.
XX
OS Chimeric - Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1..2004
FT /tag= a
FT /product= "VP22-Cre fusion protein"
XX
XX WO200149837-A2.
XX
XX 12-JUL-2001.
XX
XX 05-JAN-2001; 2001WO-EP00060.
XX
XX 07-JAN-2000; 2000EP-0100351.
XX
XX 10-NOV-2000; 2000EP-0124595.
XX
XX (ARTE-) ARTEMIS PHARM GMBH.
XX
XX Schwenk F;
XX
XX PI
XX
XX WPI; 2001-441873/47.

```

```

DR P-PSDB; AAE05266.
XX
XX Using site-specific DNA recombinase domain/protein transduction domain
XX fusion proteins for inducing target gene alterations in organisms or
XX cell cultures.
XX
XX Claim 17; Page 32-35; 85pp; English.
XX
XX The present invention relates to use of fusion proteins comprising
XX a site-specific DNA recombinase domain e.g. Cre and a protein
XX transduction domain (PTD) e.g. the Human immunodeficiency virus
XX (HIV) derived Tat peptide, for preparing an agent for inducing
XX target gene alterations in a living organism or cell culture. The
XX present invention also provides a method for inducing gene
XX alterations in living organisms using the fusion proteins of the
XX invention. The present sequence is a DNA encoding VP22-Cre fusion
XX protein. The VP22 sequence is from Human spumaretrovirus (HSV).
XX
SQ Sequence 2004 BP; 419 A; 608 C; 592 G; 385 T; 0 other;

alignment_scores:
Quality: 1759.00 Length: 344
Ratio: 5.113 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.419

alignment_block:
US-09-662-128A-2 x AAD09259 ..

Align seg 1/1 to: AAD09259 from: 1 to: 2004

7 LysValSerAsnLeuLeuThrValHisGlnAsnLeuProAlaLeuProVa 23
970 GCGATGTCGAATTACTGACCGTACACCAAAATTTGCCATCTACCGGT 1019
23 LAspAlaThrSerAspGluValAlaArgLysAsnLeuMetAspMetPheArg 40
1020 CGATCAACGAGATGATGAGTTCGCCAAGAACCTGATGACATGATGAGG 1069
40 SPArgGlnAlaPheSerGluHisThrTrpLysMetLeuLeuSerValCys 56
1070 ATCCGCGAGCGCTTTCTGAGCATACCTCGGAATAATGCTTGTCCGTTGG 1119
57 ArgSerTrpAlaAlaTrpCysLysLeuAsnAsnArgLysTrpPheProAl 73
1120 CGGTGCGTGGCGCGCATGCTCAAGTTGAATACCGGAATGTTTCCGCC 1169
73 aGluProGluAspValArgAspTrpLeuLeuTrpLeuGlnAlaArgGlyL 90
1170 AGAACCTGAAGATGTTCCGATTAATCTTCTATATCTTCAGGCGCGGCTC 1219
90 eAlaValLysTrpTrpLleGlnHisLeuGlyGlnLeuAsnMetLeuHis 106
1220 TGGCAGTAAAACTATCCAGCAACATTTTGGCCAGCTAAACATGCTTCAT 1269
107 ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuVa 123
1270 CGTGGTCCGGCTGCCACGACCAAGTGAACGCAATGCTTTCACCTGCT 1319
123 lMetArgArgLleArgLysGluAsnValAspAlaGlyLleuArgAlaLysG 140
1320 TATGGCGCGATCCGAAGAAAACGTTGATCCGCTGACCTGCAACAAAC 1369
1320 lAlaLeuAlaPheGluArgTrpAspPheAspGlnValArgSerLeuMet 156
1370 AGGCTCTAGCGTTGCAACGCACTGATTTGACCAAGCTTCCTCACTCATG 1419
157 GluAsnSerAspArgCysGlnAspLleArgAsnLeuAlaPheLeuGlyL1 173
1420 GAATAATGCGATGCTGCCAGATATACGTAATCGCATTTCTGGGGAT 1469
173 eAlaTrpAsnThrLeuValArgLleAlaGluLleAlaArgLleArgValL 190

```

```

1470 TGCTTAACACCCCTGTACGTATAGCCGAATTCACAGATCAGGTTA 1519
190 ysaapilesararqthraspglyargmetleuilehisileglyarg 206
1520 AAGATATCTCAGTACTAGCGGGAGAAATTTAATCATTTGGCAGA 1569
207 ThrLysThrLeuValSerThrAlaGlyAlaGluLysAlaLeuSerLeu 223
1570 ACGAAACGGTGGTTAGACCCGAGGTGTAGAAAGCACTTAGCCTGG 1619
223 yvalThrLysLeuValGluArgTrpIleSerValSerGlyValAlaAsp 240
1620 GGTAACATAACCTGTCGAGGATGATTCCTGCTGGTGTAGCTGATG 1669
240 sspRoAsnArgLysLeuPheCysArgValArgLysAsnGlyValAlaAla 256
1670 ATCCGAATTAACCTACTCTTTTGGCGGGTCAGAAAAATGGTGTGG 1719
257 ProSerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyLeuPheG 273
1720 CCATCTGCACACGACGCTATCACTCGCGCCCTGGAAAGGATTTTGA 1769
273 ualAThHisArgLeuIleTyrGlyAlaLysAspSerGlyGlnArgT 290
1770 AGCAACTCATGATTGATTACGGCGCTAAGCATGACTGTGTCAGAGAT 1819
290 yLeuAlaTrpSerGlyHisSerAlaArgValGlyAlaAlaArgAspMet 306
1820 ACCTGGCTGGTGTGACACAGTCCCGTGTGGAGCCGGCGCAGATATG 1869
307 AlaArgAlaGlyValSerIleProGluIleMetGlnAlaGlyLeuTrp 323
1870 GCCCGCGTGGAGTTTCATACCGGAGATCATGCAAGCTGTGGCTGAC 1919
323 rAsnValAsnIleValMetAsnTyrIleArgAsnLeuAspSerGluThr 340
1920 CATGTAAATATTTGTCATGAACATAATCCGTAACCTGGATGTGAAACAG 1969
340 lYAlaMetValArgLeuLeuGluAspGlyAsp 350
1970 GGGCAATGGTGGCTGCTGGAGAAATGGCGAT 2001

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAD09268
seq_documentation_block:
ID AAD09268 standard; DNA: 2055 BP.
XX
AC AAD09268;
XX
DT 12-SEP-2001 (first entry)
XX
DE VP22Crestreptag fusion DNA.
XX
KW DNA recombinase domain; protein transduction domain; PTD;
KW VP22Crestreptag fusion protein; Human immunodeficiency virus; HIV;
KW gene alteration; Human spumaretrovirus; HSV; ds.
XX
OS Chimeric - Human spumaretrovirus.
OS Chimeric - unidentified.
XX
FH key 1..2052 location/qualifiers
FT CDS /*tag= a
FT /product= "HSV VP22Crestreptag fusion protein"
XX
XX WO200149832-A2.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-EP00060.
XX
PR 07-JAN-2000; 2000EP-0100351.
PR 10-NOV-2000; 2000EP-0124595.

```

```

XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Schwenk F;
XX
DR WPI; 2001-441873/47.
XX
DR P-PSDB; AAE05273.
XX
PT Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PT cell cultures -
XX
PS Claim 17; Page 55-58; 85pp; English.
XX
CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived Tat peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a DNA encoding VP22Crestreptag
CC fusion protein. The VP22 sequence is from Human spumaretrovirus
CC (HSV).
XX
SQ Sequence 2055 BP; 427 A; 619 C; 609 G; 400 T; 0 other;

alignment_scores:
Quality: 1759.00 Length: 344
Ratio: 5.113 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.419

alignment_block:
US-09-662-128A-2 x AAD09268 ..
Align seg 1/1 to: AAD09268 from: 1 to: 2055

7 LysValSerAsnLeuLeuThrValHisGlnAsnLeuProAlaLeuProAla 23
:.....:
970 CGCATGTCCTCAATTACTGACCGTACCCACCAAAATTTGCTGCATTACCGGT 1019
23 lAspAlaThrSerAspGluValArgLysAsnLeuMetAspMetPheArg 40
:.....:
1020 CGATGCAACGAGTGTAGTGGTTCGCAAGAACCTGATGACATGTTCAGGG 1069
40 sPArgGlnAlaPheSerGluHisThrTrpLysMetLeuSerValCys 56
:.....:
1070 ATCGCCAGCGCTTTCTGACATACCTGGAATAAGCTTCTGTCGCTTGC 1119
57 ArgSerTrpAlaIleTrpCysLysLeuAsnAsnArgLysTrpPheProAl 73
:.....:
1120 CGGTGCTGGGGCGCATGGTGCAGATGTAATACCGAAATGGTTCCCGC 1169
73 agLupProGluAspValArgAspTyrLeuLeuTyrLeuGlnAlaArgGly 90
:.....:
1170 AGAACCTGAAGATGTTCCGATTAATCTTATATCTTAGCGGGCGGCT 1219
90 euAlaValLysThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHis 106
:.....:
1220 TGGCAGTAATAAACTATCCAGCAACATTTGGCGCACCTAAACATGGTTCAT 1269
107 ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuVal 123
:.....:
1270 CGTCGGTCCGGGCTGCCACAGCAAGTGCACCAATGCTGTTCACGTGT 1319
123 lMetArgArgIleArgLysGluAsnValAspAlaGlyGluArgAlaLysG 140
:.....:
1320 TATCGCGCGGATCCGAAGAAACGTTGATGCCGCTGACACTGCAAAAC 1369
140 lAlaLeuAlaPheGluArgThrAspPheAspGlnValArgSerLeuMet 156
:.....:
1370 AGGCTCTAGCGTTGCAAGCAGCTGATTCGACAGGTTGTTCACTCATAG 1419

```

```

157  GluAsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyI 173
1420  GAAATATAGGATCGCTGCCAGATATACCTAATCTGCGATTCTTGCGGAT 1469
173  eAlaIYrAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArgValL 190
1470  TGGTATATACACCCCTGTACGTATAGCCGAATATCCAGATACAGGGTTA 1519
190  yAspIleSerArgThrAspGlyArgMetLeuIleHisIleGlyArg 206
1520  AGATATCTCAGCTACTGACGGTGGGAGATGTATATCCATATTGGCGCA 1569
207  ThrIYsThrLeuValSerThrAlaGlyValGluYsAlaLeuSerLeuG 223
1570  ACGAAATACCGCTGTAGCACCCAGCTAGAGAAAGCACTTACCTCGG 1619
223  yValThrIYsLeuValGluArgTrpIleSerValSerGlyValAlaAsp 240
1620  GGTACTATAACTGCTGACGCGATGATTCCTCTCTGCTGTACTGATG 1669
240  sPProAsnAsnIYrLeuPheCysArgValArgIYsAsnGlyValAla 256
1670  ATCCGAATTAACCTACTGTTTCCGGGCTCAGAAAAAATGGTGTCCGCG 1719
257  ProSerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyIlePheG 273
1720  CCATCTGCCACAGCCAGCTATCACTCGCCCTCGAAGAGATTTTGA 1769
273  uAlaThrIAsArgLeuIleYrGlyAlaIYsAspSerGlyGlnArgT 290
1770  AGCAACTCATGATGATTACGGCGCTAAGATGACTCTGCTCAGAGAT 1819
290  YrIleuAlaTrpSerGlyHisSerAlaArgValGlyAlaIAspMet 306
1820  ACCTGGCTGCTGACACAGTCCCGCTCGGAGCCGCGCAGATATG 1869
307  AlaArgAlaGlyValSerIleProGluIleMetGlnAlaGlyIYrPhe 323
1870  GCCCGCGCTGAGTTTCATATCCGAGATCATGCAAGCTGTGGTGAGC 1919
323  rAsnValAsnIleValMetAsnTYrIleArgAsnLeuAspSerGlyThrG 340
1920  CAAATGAAATGTGTCAATGAACTATATCCGTAACTGGATGAGAAACAG 1969
340  IYAlaMetValArgLeuLeuGluAspGlyAsp 350
1970  GGGCAATGTGCGCTGCTGGAAGATGGCGAT 2001

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAD09266
seq_documentation_block:
ID  AAD09266 standard: DNA: 4488 BP.
XX
AC  AAD09266;
XX
DT  12-SEP-2001 (first entry)
XX
DE  PT7-VPCS vector DNA related to the invention.
XX
KW  DNA recombinase domain; protein transduction domain; PTD;
KW  gene alteration; fusion protein; Human immunodeficiency virus;
KW  HIV; PT7-VPCS vector; ds.
XX
OS  Unidentified.
XX
PN  WO200149832-A2.
XX
PD  12-JUL-2001.
XX
PF  05-JAN-2001; 2001WO-EP00060.
XX
PR  07-JAN-2000; 2000EP-0100351.

```

```

PR  10-NOV-2000; 2000EP-0124595.
XX
XX  (ARTE-) ARTEMIS PHARM GMBH.
XX
PI  Schwenk F.
XX
DR  WPI: 2001-441873/47.
XX
PT  Using site-specific DNA recombinase domain/protein transduction domain
PT  fusion proteins for inducing target gene alterations in organisms or
XX  cell cultures -
XX
PS  Example; Page 51-52; 85pp; English.
XX
CC  The present invention relates to use of fusion proteins comprising
CC  a site-specific DNA recombinase domain e.g. Cre and a protein
CC  transduction domain (PTD) e.g. the Human immunodeficiency virus
CC  (HIV) derived TAR peptide, for preparing an agent for inducing
CC  target gene alterations in a living organism or cell culture. The
CC  present invention also provides a method for inducing gene
CC  alterations in living organisms using the fusion proteins of the
CC  invention. The present sequence is a PT7-VPCS vector DNA related to
XX  the invention.
XX
SQ  Sequence 4488 BP; 1038 A; 1239 C; 1190 G; 1021 T; 0 other:

alignment_scores:
      Quality: 1759.00      Length: 344
      Ratio: 5.113          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 99.419

alignment_block:
US-09-662-128A-2 x AAD09266 ..

Align seg 1/1 to: AAD09266 from: 1 to: 4488

7  LysValSerAsnLeuLeuThrValHisGlnAsnLeuProAlaLeuProAla 23
   ::::|||||
3369  CGCATGTCCAAATTACTGACCGCTACACCAAAATTTGCTGCTCATACCGGT 3418
23  lAspAlaThrSerAspGluValArgIYsAsnLeuMetAspMetPheArgA 40
   |||||
3419  CGATGCAACAGATGATAGGTGTGCAAGAACCTGATGGACATGTTCAGGG 3468
40  sPArgGlnAlaPheSerGluHisThrTrpIYsMetLeuSerValYs 56
   |||||
3469  ATGCCAGGCGTTTCTGACCATACCTGGAATAATGCTTGTCTGCTCCGTTC 3518
57  ArgSerTrpAlaAlaTrpCysIYsLeuAsnAsnArgLYsTrpPheProAl 73
   :|||
3519  CGGTGCGGGCGGCATGTGTCAGATGTAACCGAAATGTTTCCCGC 3568
73  aGluProGluAspValArgAspTYrLeuLeuTYrLeuGlnAlaArgGlyL 90
   |||||
3569  AGAACCTGAAGATGTCCGATATATCTTATATGCTTACGGCGCGGTC 3618
90  eValAlaValYsThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHis 106
   |||||
3619  TGGCAGTAAATATCATCCAGCAATTTGGCCACCTAAACATGCTTCAT 3668
107  ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuVal 123
   |||||
3669  CGTGCTCCGGGCTGCCACGACCAAGTGCACACATGCTGTTTCACTGCT 3718
123  lMetArgGlyIleArgLYsGluAsnValAspAlaGlyValArgAlaYsG 140
   |||||
3719  TATGCGCGGAGATCCGAAAGAAACGTTGATGCCGTGAACGTCCAAAC 3768
140  lAlaLeuAlaPheGluArgThrAspPheAspGlnValArgSerLeuMet 156
   |||||
3769  AGGCTCTACGTTTCGACGCACTGATTTTCAGACAGGTTCTTCACTCATG 3818

```

```

157  GUAASerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGly11 173
      |||||||
3819  GAAATATGCGATCGCTGCCAGATATACGTAATCTGGCATTTCTGGGAT 3868
      |||||||
173  eAlaTyrAsnThrLeuAlaArgIleAlaGlyIleArgVal1 190
      |||||||
3869  TGGTTATTAACCCCTGTACGTATAGCCGAAATGGCCAGATAGGCTTA 3918
      |||||||
190  yAspIleSerArgThrAspGlyGlyIleArgMetLeuIleHisIleGlyArg 206
      |||||||
3919  AAGATATCTCAGTCTACGCTGCGTGGAGATGTTAATCATATTTGGAGA 3968
      |||||||
207  ThrIleThrLeuValSerThrAlaGlyAlaGlyIleValLeuSerLeuG1 223
      |||||||
3969  ACGAAATACGCTGTAGACCGCAGGTGTAGAGAGGACCTTAGCCCTGGG 4018
      |||||||
223  yAlaThrIleValAlaGlyArgTrpIleSerValSerGlyValAlaAspA 240
      |||||||
4019  GGTAACTAACTGGTCGAGCGATGGATTTCCGCTCTGCTGGTGTAGCTATG 4068
      |||||||
240  sPProAsnAsnTyrLeuPheCysArgValAlaArgIleAsnGlyValAla1 256
      |||||||
4069  ATCGAATTAACCTACTGTTTGGCCGGGTACGAAATATGTTGGCCGCG 4118
      |||||||
257  ProSerAlaThrSerGlnLeuSerThrArgAlaLeuGlnGlyIlePheG1 273
      |||||||
4119  CCATGTGCCACCGACGACCTATCAACCTGCCCGCTGGAGAGGATTTTGA 4168
      |||||||
273  uAlaThrHisArgLeuIleTyrGlyAlaIleAspAspSerGlyIleArgT 290
      |||||||
4169  ACGCACTCATGATGATTGATTCGGGGCTTAAGATGACTCTGTGCAGAGAT 4218
      |||||||
290  yLeuAlaThrSerGlyHisSerAlaArgValAlaAlaAlaArgAspMet 306
      |||||||
4219  ACCGTGCCCTGGTCTGGACACATGCGCTGTGGAGCGCGCGAGATATG 4268
      |||||||
307  AlaArgAlaGlyValSerIleProGluIleMetGlnAlaGlyIleTrpTh 323
      |||||||
4269  GCCCGCGCTGGAGTTCATACCGAGATACGCAAGCTGGTGGCTGAGC 4318
      |||||||
323  rAsnValAsnIleValMetAsnTyrIleArgAsnLeuAspSerGlyIleTrp 340
      |||||||
4319  CAATGTAAATATGTCATGACTATATCCGTAACTGGATGATGTAACAG 4368
      |||||||
340  lValMetValArgLeuLeuGlnAspGlyAsp 350
      |||||||
4369  GGGCAATGGTGGCTGCTGGAAGATGGCGAT 4400
      |||||||
seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAc62541
seq_documentation_block:
ID AAC62541 standard; DNA: 1172 BP.
XX
AC AAC62541;
XX
DT 07-FEB-2001 (first entry)
XX
DE Cre wild-type coding sequence.
XX
KW Cre variant recognition site; lox site; recombinase;
KW variant recombination site; hybrid crop production; seedless crop;
KW phage packaging; cloning; ds.
XX
OS Unidentified.
XX
PN WO200060091-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09154.
XX
PR 06-APR-1999; 99US-0127977.
XX

```

```

PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Sauer BL, Rufer AW;
XX
DR WPI; 2000-665010/64.
XX
PT Identifying variant recombinases mediating recombination at variant
PT sites (VRS) by contacting a mutant recombinase, a first and second VRS
PT having a reporter gene, and a second nucleic acid having 2 VRS and a
PT reporter gene -
XX
PS Example 1; Page 98-101; 144pp; English.
XX
CC The present invention relates to the identification of recombinase
CC variants which have an altered specificity. They are tested using
CC constructs containing variant recognition sites, which are not recognised
CC by non-mutant recombinase but undergo recombination in the presence of a
CC variant enzyme. Variant recombinases are useful in the production of
CC genetically modified crop plants, particularly seedless varieties, and in
CC phage packaging, which has uses in cloning.
XX
SQ Sequence 1172 BP; 299 A; 263 C; 323 G; 287 T; 0 other;

```

```

alignment_scores:
Quality: 1758.00      Length: 344
Ratio: 5.110          Gaps: 0
Percent Similarity: 100.000      Percent Identity: 99.419

```

```

alignment_block:
US-09-662-128a-2 x AAC62541 ..

```

```

Align seg 1/1 to: AAC62541 from: 1 to: 1172

```

```

7  LysValSerAsnLeuLeuThrValHisGlnAsnLeuProAlaLeuProVa 23
      |||||||
53  GAAATGTCATTTCTGACCGGTACACCAAAATTTGCTGCATTAACCGGT 102
      |||||||
23  lAspAlaThrSerAspGluValAlaArgIleAsnLeuMetAspMetPheArg 40
      |||||||
103  CGATGCACAGAGTATGAGTTCGCCAAGACCTGATGACATGTCAGGG 152
      |||||||
40  sPArgGlnAlaPheSerGlnHisThrTrpIlePheMetLeuSerValCys 56
      |||||||
153  ATGCGCAGCGCTTTTCTGACATTAACCTGGAATTAAGCTTCTGCGTTGC 202
      |||||||
57  ArgSerTrpAlaAlaTrpCysIleAsnAsnArgIleTrpPheProAl 73
      |||||||
203  CGGTGCTGGCGGCGCATGTCGCAAGTTGAATACCGGAATGTTCCGCG 252
      |||||||
73  aGluProGluAspValArgAspTyrLeuLeuTyrLeuGlnAlaArgGlyL 90
      |||||||
253  AGAACCTGAAGATGTCGCGATTAATCTTATATTTCAAGCGCGCGGTC 302
      |||||||
90  eAlaValIleYThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHis 106
      |||||||
303  TGGCAGTAAATACTATCCAGCAATTTGGCCACCTAAACATGCTTCAT 352
      |||||||
107  ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuVa 123
      |||||||
353  CGTCCGTCCGCGCTGCCACGACCAAGTACACCAATGCTGTTCACTGCT 402
      |||||||
123  lMetArgArgIleArgIleGlnAsnValAspAlaGlyIleArgAlaValys 140
      |||||||
403  TATGCGCGGATCCGAAAGAAACCTTATGCCGCTGACAGCTGCAAAAC 452
      |||||||
140  lAlaLeuAlaPheGluArgTrpAspPheAspGlnValArgSerLeuMet 156
      |||||||
453  AGGCTTAGCGTTTCGAACGACATGATTCGACAGAGTTGTTCACTCATG 502
      |||||||
157  GUAASerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGly11 173
      |||||||
503  GAAATATGCGATCGCTGCCAGATATACGTAATCTGGCATTTCTGGGAT 552
      |||||||

```

173 eAlaTyrAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArgValL 190
|||||
553 TGCTTATAACACCCCTGTACGTATAGCCGAATTCAGAGATCAGGCTTA 602
190 yAspIleSerArgThrAspGlyGlyArgMetLeuIleHisIleGlyArg 206
|||||
603 AAGATTACTCTACGTACTGACGTGGAGAGATTAAATCCATATTGGCAGA 652
207 ThrLysThrLeuValSerThrAlaGlyValGluLysAlaLeuSerLeuG 223
|||||
653 ACGAAAACGCTGGTACACCCGAGCTGTAGACAGAGCACTTAGCCTGGG 702
223 yValThrLysLeuValGluArgTrpIleSerValSerGlyValAlaAspA 240
|||||
703 GGTAACTAACTGCTGACGAGATGATTCCTGCTCTGGGTAGCTGATG 752
240 sProAsnAsnTyrLeuPheGlyArgValArgLysAsnGlyValAlaAla 256
|||||
753 ATCCGAATTACTACTCTGTTTGGCCGGTCAGAAAAAATGTGTGCCGCG 802
257 ProSerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyIlePheG 273
|||||
803 CCATCTGCCACCAAGCAGCTATCACTGCCGCCCTGGAGAGGATTTTGA 852
273 uAlaThrHisArgLeuIleTyrGlyAlaLysAspAspSerGlyGlnArgT 290
|||||
853 AGCAACTCATGATGATTACGGCGCTAAGAGTAGCTCTGCTCAGAGAT 902
290 yrLeuAlaTrpSerGlyHisSerAlaArgValGlyAlaAlaArgAspMet 306
|||||
903 ACCTGGCCGTGCTGTGACACAGTGCCTGCTGGAGCCGCCGAGATATG 952
307 AlaArgAlaGlyValSerIleProGluIleMetGlnAlaGlyIleTrpTh 323
|||||
953 GCCCGCGGTGAGTTCAATACCGAGATCATGCAAGCTGTGGCTGGAC 1002
323 rAsnValAsnIleValMetAsnTyrIleArgAsnLeuAspSerGluThrG 340
|||||
1003 CAATGTAAATATTGTCAATGAATATATCCGTAACTGTGATAGTGAACAG 1052
340 yAlaMetValArgLeuLeuGluAspGlyAsp 350
|||||
1053 GGGCAATGTGTGGCCTGCTGGAAGATGGCGAT 1084

OM of: US-09-662-128a-2 to: Issued_Patents_NA: * out_format : pfs
Date: May 10, 2002 4:10 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=firmet.p2n.model -DEV=xip
-O=/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-457-342-1 + 86.00 126.02 24.86
-DB=Issued_Patents_NA -QPM=fastap -SUFFIX=tbl -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsun62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pcp
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPR=pfis
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09662128.@cgn1_1.85 -NCP=6 -ICPU=3 -LONGLOG
-DEV_TTIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-662-128a-2
Query length: 350
Database: Issued_Patents_NA: *
Database sequences: 38353
Database length: 122816752
Search time (sec): 71.710000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-193-503B-2		1757.00	3674.50	5.5e-197	1032	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-08-864-224-10		1757.00	3668.04	1.3e-196	1740	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-193-503B-5		1757.00	3664.34	2.0e-196	2346	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-193-503B-4		1754.00	3658.01	4.6e-196	2346	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-193-503B-8		1754.00	3658.01	4.6e-196	2346	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-08-735-609-11		671.00	1388.52	1.2e-69	699	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-315-372-11		671.00	1388.52	1.2e-69	699	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-244-752-11		671.00	1388.52	1.2e-69	699	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-245-497-11		671.00	1388.52	1.2e-69	699	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-105-390-51		103.50	186.91	0.0101	1008	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-125-645-11		96.50	127.26	21.21	37948	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-096-867-1		95.50	135.78	7.11	16075	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-08-814-052-12		92.00	159.62	0.3340	1288	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-08-743-637B-171		92.00	159.62	0.3340	1288	1	
/cgn2_6/prodata/2/lna/5B_COMB.seq:US-08-526-840B-171		92.00	159.62	0.3340	1288	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-09-028-934-28		91.50	134.61	8.26	8931	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-764-233A-1		91.50	113.45	1.4e-05	4403765	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-09-103-840A-2		90.00	151.62	0.9324	1749	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-116-098-1		89.00	142.62	2.36	3619	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-09-105-390-2		89.00	142.62	2.36	3619	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-09-103-840A-2		89.00	142.62	2.36	3619	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-09-103-840A-1		89.00	142.62	2.36	3619	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-481-206-13		88.00	139.38	4.48	3344	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-481-206-13		88.00	139.38	4.48	3344	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-09-103-840A-1		87.50	49.43	3.4e-05	4411529	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-09-103-840A-1		86.00	126.02	24.86	7001	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-258-261B-1		86.00	126.02	24.86	7001	1	
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-456-837-1		86.00	126.02	24.86	7001	1	

/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-457-342-1 + 86.00 126.02 24.86
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-457-646A-1 + 86.00 126.02 24.86
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-458-076A-1 + 86.00 126.02 24.86
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-457-335A-1 + 86.00 126.02 24.86
/cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-761-258-6 + 86.00 126.02 24.86
seq_name: /cgn2_6/prodata/2/lna/5B_COMB.seq:US-09-193-503B-2

seq_documentation_block:

Sequence 2, Application US/09193503B
Patent No. 6262341
GENERAL INFORMATION:
APPLICANT: Basczyński, Christopher L.
APPLICANT: Lysnik, Leszek A.
APPLICANT: Gordon-Kamm, William J.
APPLICANT: Guan, Xueni
APPLICANT: Rao, Guru
APPLICANT: Taglianti, Laura A.
TITLE OF INVENTION: A No. 6262341 Method For The Integration Of Foreign DNA Int
FILE REFERENCE: 5718-66 (amended listing)
CURRENT APPLICATION NUMBER: US/09193,503B
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/099,435
PRIOR FILING DATE: 1998-09-08
PRIOR APPLICATION NUMBER: 60/056,627
PRIOR FILING DATE: 1997-11-18
PRIOR APPLICATION NUMBER: 60/065,613
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Nucleotide
OTHER INFORMATION: sequence encoding Cre protein from Bacteriophage
US-09-193-503B-2

alignment_scores:

Quality: 1757.00	Length: 343
Ratio: 5.122	Gaps: 0
Percent Similarity: 100.000	Percent Identity: 99.708

alignment_block:

US-09-662-128a-2 x US-09-193-503B-2 ..

Align seg 1/1 to: US-09-193-503B-2 from: 1 to: 1032

8 ValSerAsnLeuThrValHISGlnAsnLeuProAlaLeuProValAs 24
1 ATGTCCACCTGCTCAGGTCACGACGACCTTCGCGCTTCAGTGA 50
24 palatirSerAspGluValArgLysAsnLeuMetLysMetPheArgAspA 41
51 CGGAGCTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 100
41 rglAlaPheSerGluHisThrTrpLysMetLeuSerValCysArg 57
101 GGGAAACGCTTACGCGACGACGACGACGACGACGACGACGACGACG 150
58 SerTrpAlaAlaTrpCysLysLeuAsnAsnArgLysTrpPheProAlaGl 74
151 TCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 200
74 uProGluAspValArgAspTrpLeuLeuTrpLeuGlnAlaArgGlyLeuA 91
201 GCCCGAGCAGCTGAGGATTAACCTTCTGACTGCAACCTGCGGCGCTGG 250
91 laValLysThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHisArg 107


```

964 AAAACATCCAGACATTTGGCCAGCTAAACATGCTTCATCGTGGTC 1013
109 TGLTLEPRARIGPROSERASPSERASNALVALSERLEUVALMETARGA 126
1014 CCGGCTGCCAGACAGTACAGCAATGCTGTTCCACGTGTTATGCGGC 1063
126 TGLTLEARGLSGLUASNVALASPALAGLYLUARGALATSGINLAEU 142
1064 GGAATCCGAAAGAAAGCTTGATGCCGGGGAACGTCGCAAAACGCTCTA 1113
143 ALAPHGLUARGTHRASPHEASPGINVALARGSERLEUMETGLUASNE 159
1114 GCGTTGCAAGCAGCTGATTTGACACCGAGTTCCTCATCATGGAATAATG 1163
159 RASPARCYSGINASPILEARGASNLEUALPHLEUGLYLEALATYRA 176
1164 CGATCGCTCCAGGATATACGTAATCTGGCATTTCTGGGATGCTTATTA 1213
176 snthrleuuarlyleaglulilealarglyleargvallylsaspile 192
1214 ACACCTGTTAGCTATAGCCGAAATTCGACGATCAGGCTTAAGATATC 1263
193 SERARGTHRASPGLYGLYARGMETLEULIENHISLEGYARGTHLYSTH 209
1264 TCACGTAAGTACGCGTGGAGAAATGTTAAATCCATATTGGCAGAACGAAAC 1313
209 rleuvalserthrallaglyvalglulysalaleuenserleuglyvalthrl 226
1314 GCTGGTTACACCGCAGGTGTAGAGAACGACTAGCCTGGGGGTAACTA 1363
226 YLEUVALGLUARGTRPLIESERVALSERGLYVALALASPPROASN 242
1364 AACTGTCGAGCGATGATTCCTGCTGCTGTAGCTGATGATCCCAAT 1413
243 ASANTYRLEUPHECYARGVALARGLYSANGLYVALALAIAPROSERAL 259
1414 AACTGACTGTTTGGCGGTCAGAAAAAATGCTGTGCGCGGCACATCTGC 1463
259 ahrserglnleuSerthrargalaleuGllyllepheglnuAlaThrH 276
1464 CACGACGAGCTATCACTCGCCCTGGAGGAGGATTTTGAACCACTC 1513
276 ISARGLEULIETRYGLYVALALYSASPSPSERGLYGINARGTRYLEUALA 292
1514 ATCGATTGATTTACGGCGCTAAGATGATCTGCTCAGAGATACCTGGCC 1563
293 TRPSERGLYHISERVALARGVALGLYVALALARGASPMEALARGAL 309
1564 TGGTCTGGACACAGTCCCGCTGTGGAGCCGCGGAGATATGGCCGCC 1613
309 aGlyValSerlleProgluileMetGlnAlaglylYTrpThrAsnValA 326
1614 TGGAGTTTCATACCGAGATCATGCAAGCTGTGGCTGGACCAATGTAA 1663
326 snlllevalmetasnlyrllleargsnleuaspsergluThrGlyAlaMet 342
1664 AATATTGCATGAACCTATATCCGTAACTGATAGTGAACAGGGCAATG 1713
343 ValArgLeuEnGluAspGlyAsp 350
1714 GTGCGCTGCTGGAAATGGCGAT 1737
seq_name: /cgn2_6/plodata/2/lna/6B_COMB.seq:US-09-193-503B-5
seq_documentation_block:
; Sequence 5, Application US/09193503B
; Patent No. 6262341
; GENERAL INFORMATION:
; APPLICANT: Baszczynski, Christopher L.
; APPLICANT: Lyznik, Leszek A.
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Guan, Xueni

```

```

; APPLICANT: Rao, Guru
; APPLICANT: Tagliani, Laura A.
; TITLE OF INVENTION: A No. 6262341e1 Method For The Integration Of Foreign DNA Int
; TITLE OF INVENTION: Eukaryotic Genomes
; FILE REFERENCE: 5718-66 (amended listing)
; CURRENT APPLICATION NUMBER: US/09/193,503B
; CURRENT FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/099,435
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: 60/056,627
; PRIOR FILING DATE: 1997-11-18
; PRIOR APPLICATION NUMBER: 60/065,613
; PRIOR FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 5
; LENGTH: 2346
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: sequence
; OTHER INFORMATION: encoding mCre:FlpM, Cre from Bacteriophage P1 and
; OTHER INFORMATION: FLP from Saccharomyces, both maize preferred
; NAME/KEY: CDS
; LOCATION: (1)..(2346)
US-09-193-503B-5

```

```

alignment_scores:
  Quality: 1757.00      Length: 343
  Ratio: 5.122          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 99.708

```

alignment_block:

US-09-662-128a-2 x US-09-193-503B-5 ..

Align seg 1/1 to: US-09-193-503B-5 from: 1 to: 2346

```

8 VALSERASNLEUETHRVAlHISGLINASNLEUPROALAEUPROVALAS 24
1 ARGTCACACCTGCTCAGCGTTGACACAGAACCTTCGCGCTTCACAGGA 50
24 PALATHSERASPGILUVALARGLYSASNLEUMETASPMETPHATARGASPA 41
51 CGCGACGTCGATGATGACAGAACCACTCATGACATGTCGCGACGA 100
41 RGLNALAPHSESERGLUHISThrTrpLysMetLeuSerValIcysarg 57
101 GCGAAGCGTTCAGCGACACACCTGGAAGATGCTGCTCGCTCGCGCG 150
58 SERTPRALALATPRCYSLYSLEUASNANARGLYSTRPHEPROALAGI 74
151 TCTCGGGCTGCATGCTGACAGCTGAACACAGAGAGTGTCCCGCTCA 200
74 UPROGLUASPYVALARGSPTRYRLEUETHRYRLEUGINALAARGGLYLEUA 91
201 GCCGAGGAGCGTGAGGATTAACCTTGTGTACCTGCAAGCTCGCGGGCTGG 250
91 lAValLysThrllleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHisArg 107
251 CAGTGAAGACCATTCACGACCAACCTTGGACACCTGAACATGCTTCACAG 300
108 ARGSERGLYLEUPROARGPROSERASPSERASNALVALSERLEUVALME 124
301 CGCTCCGGCTCCCGCGGCCCCACGACGACTGAAAGCCGTGATCTGTCAT 350
124 tArgArglLeArgLyGluAsnValAspAlaGlyGluArgAlaLyGlnA 141
351 GCGCGCGCATCAGGAAGAAACCTCGATGCGCGGAAAGCGCAAGGCAAGG 400
141 lAleuAlaPheGluArgThrAspPheAspGlnValArgSerLeuMetGlu 157

```

```

401 CCTCGCGTTCGAGAGACCGATTTCACCAAGCTCCGACCCTGATGAGAG 450
158 AsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyIleAl 174
|||||
451 AACACGACGAGGTGCGACGACATTAAGAACCTGGCGTTCTCCGGAATTGC 500
174 aTyraAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArgValysa 191
|||||
501 ATACACACAGCTCTCCAGATCGCGGAATTGCCGATCGCGCTGAGG 550
191 spIleSerArgThrAspGlyArgMetLeuIleHisIleGlyArgThr 207
|||||
551 ACATTAGCCGACCGACCGCGGACGAGATGCTTATCCACATTGGCAGGACC 600
208 LysThrLeuValSerThrAlaGlyValGluLysAlaLeuSerLeuGlyVa 224
|||||
601 AAGACGCTGTTTCCACCGCAGGCGTCGAAAAGCCCTCAGCCTCGGAGT 650
224 IThrIysLeuValGluArgThrIleSerValSerGlyValAlaAspAsp 241
|||||
651 GACCAAGCTGTCGAACGCTGATCTCGTGCGGCGCTCGGACGACGCC 700
241 roAsnAsnIleuPheCysArgValArgLysAsnGlyValAlaAlaPro 257
|||||
701 CAACACACTACTCTTGTGCGCGCTCCGCAAGAACGGGCTGCGCCCT 750
258 SerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyIlePheGluAl 274
|||||
751 AGGCGCACACGCAACTCAGCAGAGGCGCTGGAAAGTATTTCGAGGC 800
274 aThrHisArgLeuIleTyrglyAlaLysAspAspSerGlyAsnArgTyrL 291
|||||
801 CACCCACCGCTGATCTACGGGCGGCAAGATGACAGCGTCAACGCTACC 850
291 euAlaTrpSerGlyHisSerAlaArgValGlyAlaAlaArgAspMetAla 307
|||||
851 TCCATGATGTCGGGCGACCTCGCGCGGTGGAGTCTAGGACATGGCC 900
308 ArgAlaGlyValSerIleProGluIleMetGlnAlaGlyTyrThrAs 324
|||||
901 CGGCGCGGTTTCCATCCCGAATCATGCGAGCGGCGGATGACGAA 950
324 nValAsnIleValMetAsnTyrIleArgAsnLeuAspSerGluThrGlyAla 341
|||||
951 CGTGAACATGTCATGACTACATTCGCAACCTTGACAGCAGAGCGGCG 1000
341 laMetValArgLeuLeuGluAspGlyAsp 350
|||||
1001 CAATGTTGCTCTCTGGAAGATGGCGAT 1029
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-193-503B-4
seq_documentation_block:
; Sequence 4, Application US/09193503B
; Patent No. 6262341
; GENERAL INFORMATION:
; APPLICANT: Baszczyński, Christopher L.
; APPLICANT: Lyznik, Leszek A.
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Guan, Xueni
; APPLICANT: Rao, Guru
; APPLICANT: Tagliani, Laura A.
; TITLE OF INVENTION: A No. 6262341el Method for The Integration of Foreign DNA Into
; FILE REFERENCE: 5718-66 (amended listing)
; CURRENT APPLICATION NUMBER: US/09/193,503B
; CURRENT FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/099,435
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: 60/056,627
; PRIOR FILING DATE: 1997-11-18
; PRIOR APPLICATION NUMBER: 60/065,613
; PRIOR FILING DATE: 1997-11-18

```

```

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2346
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Nucleotide
; OTHER INFORMATION: sequence encoding a Cre:FLPm polypeptide, Cre
; OTHER INFORMATION: from Bacteriophage p1 and FLP (malize preferred
; OTHER INFORMATION: (codons) from Saccharomyces
; US-09-193-503B-4

```

```

alignment_scores:
    Quality: 1754.00      Length: 343
              Ratio: 5.114      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 99.417

```

alignment_block:

US-09-662-128A-2 x US-09-193-503B-4

Align seg 1/1 to: US-09-193-503B-4 from: 1 to: 2346

```

8 valSerAsnLeuThrValHisGlnAsnLeuProAlaLeuProValAs 24
|||||
1 ATGGCCAAATTACTGACCGCTACACCAAAATTTGCTGCATTACCGGTGCA 50
24 palArThrSerAspGluValArgLysAsnLeuMetAspMetPheArgAsp 41
|||||
51 TGCACGAGTGTAGGTTCCGAAAGAACCTGATGACATGTTACGAGGATC 100
41 rglAlaPheSerGluHisThrTrpLysMetLeuLeuSerValCysArg 57
|||||
101 GCCAGGCGTTTCTGAGCATACCTGGAATGCTTGTGCTGTTGCGCG 150
58 SerThrAlaIleTrpCysLysLeuAsnAsnArgLysTrpPheProAlaG 74
|||||
151 TCGTGGCGCGCATGTGTGCAAGTTGAATACCGGAATGTTCCGCGACA 200
74 uproGluAspValArgAspTyrLeuLeuTyrLeuGlnAlaArgGlyLeuAl 91
201 ACCTGAAGATGTTGCGGATTATCTTATATCTTACGCGCGCGGTGCG 250
91 laValIysThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHisArg 107
|||||
251 CAGTAAAACTATCCAGCAACATTTGGCCAGCTTAAATGCTTCAATCGT 300
108 ArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuValMe 124
|||||
301 CGGTCCGGGCTGCGACAGCAAGTGACACCAATGCTGTTCACTGTTAT 350
124 tArgArgIleArgLysGluAsnValAspAlaGlyGluArgAlaLysGlnAl 141
351 GCGGCGGATCCGAAAGAAACGTTGATCGGTGAACGTCGAAACACGAG 400
141 laLeuAlaPheGluArgThrAspPheAspGlnValArgSerLeuMetGlu 157
|||||
401 CTCTAGGTTTCAAGCAGCATGTTGACAGAGTTGCTTCACTCATGTGAA 450
158 AsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyIleAl 174
451 AATAGCATGCTGCCAGATATACGTAATCTGGCATTTCTGCGGATATGC 500
174 aTyraAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArgValysa 191
|||||
501 TTATTAACACCTGTTACGATATACCGAAATTTGCCAGAGATCAGGTTAAG 550
191 spIleSerArgThrAspGlyArgMetLeuIleHisIleGlyArgThr 207
|||||
551 ATATCTCAGTACGACGAGTGGGAGATGTTATTCATATATGGCAGAAACG 600
208 LysThrLeuValSerThrAlaGlyValGluLysAlaLeuSerLeuGlyVa 224

```

```

601 AAAACGCTGGTAGACACCGAGGTAGAGAAAGCACTTACCTGGGGGT 650
224 LThrLysLeuValGluArgTrpLysSerValSerGlyValAlaAspAsp 241
651 AACTAAACCTGGTCAGCGAGATGATTCCTCTCTGCTGTAGCTGATGATC 700
241 rosnasnTrpLeuPheCysArgValArgLysasnGlyValAlaAlaPro 257
701 CGAATAACCTGCTGTTTGGCGGGGTACGAAAAATGGTGTCCCGGCCA 750
258 SerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyLeuPheGluAl 274
751 TGTGCCACACGCGCAGCTATCACTCCGCCCTGGAAAGGATTTTGAAGC 800
274 aThrHisArgLeuIleTyrGlyAlaLysAspSerGlyGlnArgTyrL 291
801 AACTCATGATGATTTACGGCGCTAAGATGATCTGTGTCAAGATATCC 850
291 euAlaTrpSerGlyHisSerAlaArgValGlyAlaAlaArgAspMetAla 307
851 TGGCTGGTGTGGACACACTGCCCTGTGCGAGCGCGCGAGATATGGCC 900
308 ArgAlaGlyValSerIleProGluIleMetGlnAlaGlyIleTyrPheAs 324
901 CGGGCTGGAGTTTCAATACCGAGATCATGCAAGCTGGCTGGACCAA 950
324 nValAsnIleValMetAsnTyrIleArgAsnLeuAspSerGlyIleThrGly 341
951 TGTAAATATGTCATGAACTATATCCGTACCTGGATGTGAAACACAGGG 1000
341 laMetValArgLeuLeuGluAspGlyAsp 350
1001 CAATGCTGGCCTGCTGGAAAGATGGCGAT 1029

seq_name: /cgn2_6/prodata/2/lna/6B_COMB.seq:us-09-193-503B-7

seq_documentation_block:
: Sequence 7, Application US/09193503B
: Patent No. 6262341
: GENERAL INFORMATION:
: APPLICANT: Baszczynski, Christopher L.
: APPLICANT: Lyznik, Leszek A.
: APPLICANT: Gordon-Kamm, William J.
: APPLICANT: Guan, Xueni
: APPLICANT: Rao, Guru
: TITLE OF INVENTION: A No. 6262341el Method For The Integration Of Foreign DNA Into
: FILE REFERENCE: 5718-66 (amended listing)
: CURRENT APPLICATION NUMBER: US/09/193,503B
: PRIOR APPLICATION NUMBER: 1998-11-17
: PRIOR FILING DATE: 1998-09-08
: PRIOR APPLICATION NUMBER: 60/099,435
: PRIOR FILING DATE: 1997-11-18
: PRIOR APPLICATION NUMBER: 60/065,613
: PRIOR FILING DATE: 1997-11-18
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 7
: LENGTH: 2346
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: sequence
: OTHER INFORMATION: encoding a Cre:FLP polypeptide, Cre from
: OTHER INFORMATION: Bacteriophage P1 and FLP from Saccharomyces
US-09-193-503B-7

```

alignment_scores: Quality: 1754.00 Length: 343

```

Ratio: 5.114                      Gaps: 0
Percent Similarity: 100.000                      Percent Identity: 99.417
alignment_block:
US-09-662-128a-2 x US-09-193-503B-7
Align seg 1/1 to: US-09-193-503B-7 from: 1 to: 2346

8 ValSerAsnLeuThrValHisGlnAsnLeuProAlaLeuProValAs 24
: : : : :
1 ATGGCCAAATTTACTGACCGGTACACCAAAATTTGGCTGCATTTACCGGTGA 50
24 palArthrSerAspGluValArgLysasnLeuMetAspMetPheArgAspA 41
51 TGCACAGAGATGATGAGGTTCCGACAGACCTGATGACATGTTCCAGGATC 100
41 rglInAlaPheSerGluHisThrTrpLysMetLeuLeuSerValCysArg 57
101 GCCAGGCGTTTCTGAGCATACCTGGAATAATGCTTCTGCTCCGTTGGCCG 150
58 SerTrpAlaAlaTrpCysLysLeuAsnAsnArgLysTrpPheProAlaG 74
151 TGTGGGGCGCATGGTGGCAAGTTGAATACCGGAATGGTGTCCGCGAGA 200
74 uProGluAspValArgAspTyrLeuLeuTyrLeuGlnAlaArgLysLeu 91
201 ACCGTGAAGATGTTCCGATTTATCTTCTATCTTACGAGCGCGGCTGTGG 250
91 laValLysThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHisArg 107
251 CAGTAAAAACTATCCAGACACATTTGGGCGAGCTTAAACATGCTTCAACGT 300
108 ArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuValMe 124
301 CGGTCCGGGCTGCCACGACCAAGTGACAGACCAATGCTGTTACAGGTAT 350
124 tArgArgIleArgLysGluAsnValAspAlaGlyGluArgAlaLysGln 141
351 GGGGGGAGTCCGAAAGAAAGCTTGATGCCGGTGAACGTGCAAAACAGC 400
141 laLeuAlaPheGluArgThrAspPheAspGlnValArgSerLeuMetGlu 157
401 CTCTAGCGTTTGAACGACGATGATTTGACACAGGTGCTTACTCATGAGAA 450
158 AsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyIleAl 174
451 AATAGCGATGCTGCGCAGGATATACGTAATCTGGCATTTCTGGGAGTTGC 500
174 aTyrAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArgValLys 191
501 TTATTAACACCTGTTAGCTATACCCGAAATTTGCCAAGATCAGGTTAAAG 550
191 spLieserArgThrAspGlyArgMetLeuIleHisIleGlyArgThr 207
551 ATATCTCAGCTACTGACGGTGGGAGAAATGTTATCATATTTGGCAGAACG 600
208 LysThrLeuValSerThrAlaGlyValGluLysAlaLeuSerLeuGlyVal 224
601 AAAACGCTGGTAGACACCGAGGTAGAGAAAGCACTTACCTGGGGGT 650
224 LThrLysLeuValGluArgTrpLysSerValSerGlyValAlaAspAsp 241
651 AACTAAACCTGCTGTTTGGCGGGGTACGAAAAATGGTGTCCCGGCCA 700
241 rosnasnTrpLeuPheCysArgValArgLysasnGlyValAlaAlaPro 257
701 CGAATAACCTGCTGTTTGGCGGGGTACGAAAAATGGTGTCCCGGCCA 750
258 SerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyLeuPheGluAl 274
751 TGTGCCACACGCGCAGCTATCACTCCGCCCTGGAAAGGATTTTGAAGC 800
274 aThrHisArgLeuIleTyrGlyAlaLysAspSerGlyGlnArgTyrL 291

```

```

|||||
801 AACTCATGATTGATTACGGCCCTAAGATGACTGCTGTCAGAGATACC 850
291 euA1arPserGlyHisSerAlaArgValGlyAlaAlaArgspMetAla 307
851 TGGCGTGTCTGGACACAGTGCCTGTGCGACCCGCGGAGATATGAGCC 900
308 ArgAlaGlyValSerIleProGluIleMetGlnAlaGlyIleTyrThrAs 324
901 CGCGCTGGAGTTTCAATACCGGAGATCATACGACGTGGGTGGACCAA 950
324 nValAsnIleValMetAsnTyrIleArgAsnLeuAspSerGluThrGlyA 341
951 TGTAAATATTGTGCATGAACATAATCCGTAACTGGATAGTAAGAAACAGGG 1000
341 laMetValArgLeuLeuGluAspGlyAsp 350
1001 CAATGTGCGCTGTGTAAGATGCGCAT 1029

seq_name: /cgn2_6/prodata/2/lna/6B.COMB.seq:US-09-193-503B-8

seq documentation block:
; Sequence 8, Application US/09193503B
; Patent No. 626341
; GENERAL INFORMATION:
; - APPLICANT: Baszczynski, Christopher L.
; - APPLICANT: Lyznik, Leszek A.
; - APPLICANT: Gordon-Kamm, William J.
; - APPLICANT: Guan, Xueni
; - APPLICANT: Rao, Guru
; - APPLICANT: Tagliani, Laura A.
; TITLE OF INVENTION: A No. 626341el Method For The Integration Of Foreign DNA Into
; FILE REFERENCE: 5718-66 (amended listing)
; CURRENT APPLICATION NUMBER: US/09/193,503B
; PRIOR APPLICATION NUMBER: 60/099,435
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: 60/056,627
; PRIOR FILING DATE: 1997-11-18
; PRIOR APPLICATION NUMBER: 60/065,613
; PRIOR FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2346
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: encoding a FLPM:Cpe polypeptide, FLP from
; OTHER INFORMATION: Saccharomyces (maize preferred codons), and Cpe
; OTHER INFORMATION: from Bacteriophage P1
; NAME/KEY: CDS
; LOCATION: (1)..(2346)
; US-09-193-503B-8

alignment_scores:
    Quality: 1754.00      Length: 343
    Ratio: 5.114          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 99.417

alignment_block:
US-09-662-128A-2 x US-09-193-503B-8 ..
Align seg 1/1 to: US-09-193-503B-8 from: 1 to: 2346

8 ValSerAsnLeuLeuThrValHisGlnAsnLeuProAlaLeuProValAs 24
:
1315 ATGGCAATTACTGACCGTACACCAAAATTTGCTGCTATTACCGGTGCA 1364
24 palAthrSerAspGluValAlaArgLysAsnLeuMetAspMetPheArgAspA 41

```

```

|||||
1365 TGCACAGAGTGAAGAGCTTCGACAGACCTGATGACATGTTCAAGGATC 1414
41 rgGlnAlaPheSerGlyHisThrTyrPlyMetLeuLeuSerValCysArg 57
1415 GCCAGCGGTTTCTGTGAGCATACCTGGAAATCTTCTGTGCTTTGCCGG 1464
58 SerTyrAlaAlaThrPcyLysLeuAsnAsnArgLysTyrPheProAlaG1 74
1465 TCGTGGCGGCATGTGTCAGATGATTAACCGGAATGGTTCCCGCAGA 1514
74 uproGluAspValArgAspTyrLeuLeuTyrIleuGlnAlaArgGlyLeuA 91
1515 ACCTGAAGATGTTCCGGATATTCTTATATCTTCACGGCGCGGCTGG 1564
91 laValLysThrIleGlnGlnHisLeuGlyGlnLeuAsnMetLeuHisArg 107
1565 CAGTAAAACTATCCACCAACATTGGGCCAGCTAACAATGCTTCATCGT 1614
108 ArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuValMe 124
1615 CGGTCCGGGCTGCCACAGCAAGTGAACGAAATGCTGTTCACTGGATTAT 1664
124 tArgArgIleArgLysGlnAsnValAspAlaGlyGluArgAlaLysGlnA 141
1665 GCGCGGATCCGAAAGAAACGTTGATGCCGCTGAACGTGCAAAACAGG 1714
141 laLeuAlaPheGluArgThrAspPheAspGlnValArgSerLeuMetGlu 157
1715 CTCTAGCGTTCGAACGCACTGATTCCACACAGGTGCTGTTCACTCATGA 1764
158 AsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyIleAl 174
1765 AATACCGATCGCTGCCAGATATACGTAATCTGCATTTCTGGGGAATGG 1814
174 atyAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArgValLysA 191
1815 TTATTAACACCGTGTAGCTATAGCCGAATGCCAGATCAGGGTTAAAG 1864
191 spIleSerArgThrAspGlyArgMetLeuIleHisIleGlyArgThr 207
1865 ATATCTCACGTACTGACGCTGGGAGAAATTAATCCATTATGGCAGAAC 1914
208 LysThrLeuValSerThrAlaGlyValGlnLysAlaLeuSerLeuGlyA 224
1915 AAACCGCTGTAGCACCGCAGGTAGAGAAGCACTTACCTGGGGGT 1964
224 LThrLysLeuValGluArgTyrPleSerValSerClyValAlaAspAsp 241
1965 AACTAAACTGTGTCAGCGATGATTTCCGTCTGTGTACTGATGATC 2014
241 roAsnAsnTyrIleuPheCysArgValArgLysAsnGlyValAlaAlaPro 257
2015 CGAATACTACTGTTTTCGCGGTACAGAAAATGGTGTGGCGCGCA 2064
258 SerAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyIlePheGluAl 274
2065 TCTGCCACCGACCACTATCACTCCGCCCTGGAAGGATTTTGAAGC 2114
274 atyHisArgLeuIleTyrGlyAlaLysAspAspSerGlyGlnArgTyrL 291
2115 AACTCATGATTGATTACGGCGCTAAGATGACTGTGTCAGAGATACC 2164
291 euA1arPserGlyHisSerAlaArgValGlyAlaAlaArgspMetAla 307
2165 TGGCTGTGTGGACACAGTGCCTGTGCGACCCGCGGAGATATGAGCC 2214
308 ArgAlaGlyValSerIleProGluIleMetGlnAlaGlyIleTyrThrAs 324
2215 CGCGCTGGAGTTTCAATACCGGAGATCATGCAAGCTGGTGGCTGACCAA 2264
324 nValAsnIleValMetAsnTyrIleArgAsnLeuAspSerGluThrGlyA 341

```

```
2265 TGTAAATATTGTCTGAACTATATCCGTAACTTGATGTGAAACAGGG 2314
341 laetvalargleuenglaspolyasp 350
|||||
2315 CAATGTCGCGCTGCTGGAGATGGCCAT 2343

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-735-609-11

seq_documentation_block:
; Sequence 11, Application US/08735609
; Patent No. 5953609
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Amalfitano, Andrea
; APPLICANT: Hauser, Michael A.
; APPLICANT: Kumar-Singh, Rajendra
; APPLICANT: Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,609
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; US-08-735-609-11

alignment_scores:
Quality: 671.00 Length: 127
Ratio: 5.283 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-662-128a-2 x US-08-735-609-11 ..

Align seg 1/1 to: US-08-735-609-11 from: 1 to: 699

1 MetcProLysLysLysArgLysValSerAsnLeuLeuThrValHisGlnAs 17
|||||
318 ATCCCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 367
|||||
17 nleuproAlaLeuProValAlaThrSerAspGluValArgLysAsnL 34
|||||
368 TTTCCTGCGCTTACCGCTGATGCCAACGAGTGAAGTTCGCAAGAAC 417
|||||
34 eumetAspMetPheArgAspArgGlnAlaPheSerGluHisThrTyrPly 50
|||||

2265 TGAATGACATGTTCAAGGATCCAGCGCTTTTCTGAGCATACCTGAAA 467
51 MetLeuLeuSerValCysArgSerTyrPAlaIaIaTrrCysLysLeuAsnAs 67
|||||
468 ATGCTTCTGTCCGTTTCCCGGTCGTCGGCGGCGCATGCTCAAGTGAATPA 517
|||||
67 nArgLysTrrPheProAlaGluProGluAspValArgAspTyrLeuLeuT 84
|||||
518 CCGGAAATGGTTTCCCGCAGAAACCTGAAGATGTTGGCATTAATCTTCTAT 567
|||||
84 yLeuGlnAlaArgGlyLeuAlaValLysTrrLleGlnHisLeuGly 100
|||||
568 ATCTTCAGCGCGCGGCTGTGCAGTAAACATATCCACACATTTGGGC 617
|||||
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAsp 117
|||||
618 CAGTAAACATGCTTCATCGTCGGTCCGGCTGCCACAGACCAAGTGAAG 667
|||||
117 rAsnAlaValSerLeuValMetArgArgIle 127
|||||
668 CAATGCTGTTTCACTGGTTATGCGCGCATC 698

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-735-609-11

seq_documentation_block:
; Sequence 11, Application US/08735609
; Patent No. 5994132
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Amalfitano, Andrea
; APPLICANT: Hauser, Michael A.
; APPLICANT: Kumar-Singh, Rajendra
; APPLICANT: Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,609
; FILING DATE: 23-Oct-1996
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; US-08-735-609-11

alignment_scores:
Quality: 671.00 Length: 127
Ratio: 5.283 Gaps: 0
```

ATTORNEY/AGENCY INFORMATION.
NAME: Ingolia, Diane E.

ADDRESSEE: Medlen & Carroll, LLC

STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: California
 COUNTRY: United States Of America
 ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/244,752

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/735,609

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: UM-02484

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 699 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "DNA"

US-09-244-752-11

alignment_scores:

Quality: 671.00 Length: 127
 Ratio: 5.283 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-662-128a-2 x US-09-244-752-11 ..

Align seg 1/1 to: US-09-244-752-11 from: 1 to: 699

```

1 MetProLysLysArgLysValSerAsnLeuThrValHisGlnAs 17
|||||
318 ATGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 367
17 nleuProAlaLeuProValAspAlaThrSerAspGluValArgLysAsnL 34
|||||
368 TTGGCCGTCATTCACGCGTGATGCAACGAGTATGAGTTCACCAAGAAC 417
34 eumetAspMetPheArgAspArgGlnAlaPheSerGluHisThrTrpLys 50
|||||
418 TGATGACATGTTCAAGGATCGCCAGCGCTTTCTGACGATACCTGGA 467
51 MetLeuLeuSerValCysArgSerTrpAlaAlaTrpCysLysLeuAsnAs 67
|||||
468 ATGCTTCTGTCGCTTTGCCGCGTGCGGCGCATGTCGAAGTAAATTA 517
67 nArgLysTrpPheProAlaGluProGluAspValArgAspTrpLeuLeu 84
|||||
518 CCGGAATAGTTCCTCCGAGAACCTGAACATGTCGCGATTCCTTCTAT 567
84 yrLeuGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeuGly 100
|||||
568 ATCTTCAGCGCGCGGCTGCGCAGTAAACATTCACCAACATTTGGGC 617
101 GlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSe 117
|||||
618 CAGCTAAACATGCTTCATCGTGGTCCGCGGCGCAGCAACCAAGTGAAG 667

```

117 rasnAlaValSerLeuValMetCArgArgLe 127
 |||||
 668 CATGCTGTTTCACGTCGATATCGCGCGCATC 698

seq_name: /cgn2_6/ptodata/2/1na/6a_COMB.seq:US-09-245-497-11

seq_documentation_block:

Sequence 11, Application US/09245497
 Patent No. 6083750

GENERAL INFORMATION:

APPLICANT: Chamberlain, Jeffrey S.

APPLICANT: Amalfitano, Andrea

APPLICANT: Hauser, Michael A.

APPLICANT: Kumar-Singh, Rajendra

APPLICANT: Hartigan-O'Connor, Dennis J.

TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States Of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/245,497

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/735,609

APPLICATION NUMBER: FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: UM-02484

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 699 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "DNA"

US-09-245-497-11

alignment_scores:

Quality: 671.00 Length: 127
 Ratio: 5.283 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-662-128a-2 x US-09-245-497-11 ..

Align seg 1/1 to: US-09-245-497-11 from: 1 to: 699

```

1 MetProLysLysArgLysValSerAsnLeuThrValHisGlnAs 17
|||||
318 ATGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 367
17 nleuProAlaLeuProValAspAlaThrSerAspGluValArgLysAsnL 34
|||||
368 TTGGCCGTCATTCACGCGTGATGCAACGAGTATGAGTTCGCAAGAAC 417
34 eumetAspMetPheArgAspArgGlnAlaPheSerGluHisThrTrpLys 50

```

```

|||||
418 TGATGACATGTTTCAGGATCGCAGCGGTTTCTGAGCATACCTGAA 467
51 MetleuLeuSerValCysArgSerTrpAlaIleTrpCysIleuSns 67
468 ATGCTTTCGTCGGTTCGGGTCGTGGGCGCATGGTGCAGATTGAATA 517
67 nargLysTrpPheProAlaGluProGluAspValArgAspTyrLeuLeu 84
518 CCGGAATAGCTTCCGAGACCTGAAGATGTCGCGATTATCTCTAT 567
84 TyrLeuGlnAlaArgGlyLeuAlaValAlaYstHrIleGlnIleSleuGly 100
568 ATCTTCAGCGCGCGGCTCTGGCAGTAAACATATCCAGCAACATTG6GC 617
618 CAGCTAAACATGCTTCATCTCGGTCGCGGCTGCCAGACCAAGTGACAG 667
117 AsnAlaValSerLeuValMetArgArgIle 127
668 CAATGCTGTTCACTGATTATCGCGCGATC 698

seq_name: /cgn2_6/ptodata2/ina/6B_COMB.seq:US-09-105-390-51
seq_documentation_block:
; Sequence 51, Application US/09105390
; Patent No. 6288303
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
; NUMBER OF INVENTION: and Genes
; ADDRESS: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,390
; FILING DATE: Filed herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,675
; FILING DATE: 25-JUN-97
; ATTORNEY/AGENT INFORMATION:
; NAME: Pelthory, Joanne R.
; REGISTRATION NUMBER: P42,995
; REFERENCE/DOCKET NUMBER: 2000-0455.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1008 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1008
; OTHER INFORMATION:
US-09-105-390-51

```

```

alignment_scores:
Quality: 103.50 Length: 332
Ratio: 0.719 Gaps: 16
Percent Similarity: 43.373 Percent Identity: 21.687

alignment_block:
US-09-662-128A-2 x US-09-105-390-51 ..
Align seg 1/1 to: US-09-105-390-51 from: 1 to: 1008

48 ThrTrpLysMetLeuLeuSerValCysArgSerTrpAlaIleTrpCys 64
:::|||||:
51 TCTTGAAGTTTATTCATCCCTGAGGCTGCAATCCGTTGGTGT 100
64 sLeu...AsnAsnArgLysTrpPheProAlaGluProGluAspValArg 80
|||:|||||:
101 GCTACGGCATATCGCAGACATCTCCCGCGAA..... 134
80 sPyrLeuLeuTyrLeuGlnAlaArgGlyLeuAlaValAlaYstHrIleGln 96
135 .....GAGCAGCTGCTGCAGCTCAACAAATCCAA 164
97 GlnHisLeuGlyGlnIleuAsnMetLeuHisArgArgSerGlyLeuPro 113
|||:|||||:
165 TGGCATTCACAGACATGCGCATCTTGCAGCTGAGAGCCATGACG 214
113 gProSerAspSerAsn..... 118
215 CCTCGCGGCGCAGACAGCATCGCGCTCATCGCGCTGCCCAACGACATC 264
119 .....AlaValSerLeuValMetArgArgIle 127
265 CTCATCGACCTCGCCGCCACCCGCGCTCGCGCTCGGTCAGCGC 314
128 ArgLysGluAsnValAspAlaGly.....GluArgAla 139
:::|||||:
315 GAACGTCAGCCGTTCTGCTCCGCGCGTGAACATCAAGATCTGCAGCTG 364
139 sGln..AlaLeuAlaPheGluArgThrAspPheAspGlnValArgSerLeu 155
|||:|||||:
365 GCACGAGATCTCGCGCGAGCCGACG.....CAGAACCTCTCCGCGTC 408
156 MetGluAsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeu 172
|||||:
409 ATCCAGAAC..... 417
172 YlleAlaTyrAsnThrLeuLeuArgIleAlaGluIleAlaArgIleArg 189
|||||:
418 .....ATCAAGCGCGCTGCGCGCGCGAGCATCCGCGCTCAAG 460
189 allysAspIleSerArgThrAspArgGlyArgMetLeuIleHisIle 205
:::|||||:
461 CTTCCACGCGCGGTGAAGCTAGAC.....GTC 486
206 ArgThrLysThrLeuValSerThrAlaGlyValGluYsAla...LeuSe 221
|||||:
487 GTACCAACACGTTCCCGCGCGCGCGCTGTCGCGCGCGCTCATAT 536
221 rLeuGlyValThrLysLeuValGluArgTrpLysSerValSerGlyVal 238
:::|||||:
537 GACGCGCGTGGCCACAGCTC.....CTGGATGACACGCGCGCGC 574
238 lAspAspProAsnAsnTyr...LeuPheCysArgVal...ArgLysAsn 252
|||:|||||:
575 CGCTGCTGCCAACATCTACCCCTATCTGCTACATCGCAACAAAG 624
253 GlyValAlaAlaProSerAlaThrSerGlnLeuSerThrArgAlaLeu 269
:::|||||:
625 GACATCAGCTCAACTACGACGAGTTCAGCGCGCGCGAGCTGCCGGA 674
269 u.....GlyIlePheGluAlaThrHisArgL 278
:::|||||:
675 CCCCACACGACCTGCTAGCGCAACCTGTTGAGCGCATGTGCGACT 724

```

```

278 euiletyrclalalysaspaspserglylnarгыrleuallatrpser 294
    ::::::::::: ||| ::::|
725 CCGCTACGGCGCGCTGACAAAGCCGCGC..... 753
295 G1yHisSerIalargValglYlaalAalargaspmetalaargalglYva 311
754 .....GCGGCGGCGGT 764
311 lserIleProgluIleMetGlnalaglylYtrp..... 322
    ||||| ::::: ::|||
765 CACATC.....GTGCTCGAGAGCGGCGCGCGCGCGCGGCGG 808
323 .....ThrasnValasnIleValmetasnryrIleargasnleu 335
809 ACTCGGCGCACGATCGACATCGCGCGACCTGACGACGACCTG 852
seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:us-09-251-645-11

```

seq_documentation_block:

```

: Sequence 11 Application US/09251645
: Patent No. 6281413
: GENERAL INFORMATION:
: APPLICANT: Kramer, Vance C.
: APPLICANT: Morgan, Michael K.
: APPLICANT: Anderson, Arne R.
: APPLICANT: Hart, Hope
: APPLICANT: Warren, Gregory W.
: APPLICANT: Dunn, Martha
: TITLE OF INVENTION: NOVEL INSECTICIDAL TOXINS FROM PHOTORHABDUS LUMINESCENS
: FILE REFERENCE: CGC1963/A
: CURRENT APPLICATION NUMBER: US/09/251,645
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 11
: LENGTH: 37948
: TYPE: DNA
: ORGANISM: Photorhabdus luminescens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1517)..(18035)
: OTHER INFORMATION: orf5
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (23768)..(31336)
: OTHER INFORMATION: hpn2
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (31393)..(35838)
: OTHER INFORMATION: orf2
US-09-251-645-11

```

alignment_scores:

```

Quality: 96.50 Length: 491
Ratio: 0.478 Gaps: 21
Percent Similarity: 41.141 Percent Identity: 19.756

```

alignment_block:

```
US-09-662-128a-2 x US-09-251-645-11 ..
```

```
Align seg 1/1 to: US-09-251-645-11 from: 1 to: 37948
```

```

1 MetProlyslYslYsArglyVal.....SerAsnleuLeuTh 13
   :: ||| ::||| ::|||
4579 CTAGAGAAACCAAAAGAAATTCGACGACGCCAAGAAATCTGTGAAA 4628
13 rValHISGlnAsnleuProAlaIleuProValAspAlaThrSer..... 27
   ::||| ||| |||
4629 ATTCCGGAACACGACCACTCCACAGAGACAGAAAAAAGCCGACA 4678

```

```

27 ..... 27
4679 GAAGACTGATATAAGCTGAAGACCCAGAAAAAGCCGAGAAAAACCTA 4728
28 AspGluValArglyAsnleu..... 34
   ||||| ||||| |||||
4729 GATGAGGTGAGGAAAAATCTGCCAAAAGCCATATCTCTTATATGCGCC 4778
35 .....MetaspmetrPheargAspArgGlnAlaPheSerGlnHisThrTrp. 49
   ::||| ::||| ::|||
4779 CGCTGCGCAATTAATCATCTGAAACATGCGGCAATTCCTACTCATCTGG 4828
50 .....LysMetleuLeuSerValCysArg 57
   |||||
4829 CAGATTAATCTGGAACCCGCGATAGGAAAAATGACGCGGAAAAATTTATG 4878
58 SerTrpAlaAlaIleTrpCyslyslYleuAsnAsnArglystrPrheProAlaG1 74
   ::|||
4879 AACGTGTCGGGAAATCCGTTACGGCTCAACCTGAATTCCTAATTCGACA 4928
74 uProgluAspValArgAspTyrlYleuTyrlYleuGlnAlaArgglyleuA 91
   : ||| ::|||
4929 ACCTGAACCTGTGTCCAGTAATGC..... 4953
91 lVallylThrIleGlnIleHisleuGlyGlnleuAsnMetleuHisArg 107
   ||||| ::|||
4954 .....CAGGCGCTGGCACAACTACGCGTGAATTTACCGT 4986
108 ArgserglyleuProArgProSerAspserAsnAlaValSerleu..... 122
   ::|||
4987 TCTACCGCGCTT.....ACGAAACACCTTAAGTCTGTGTGT 5024
123 .....ValMetArgglyleuArglyGluAsnValA 133
   |||||
5025 GACAAAGCCGCAACACTTGTATTAAACCAAGCACCGAAACAACCTG 5074
133 spAlaGlyGluArgAlaIleYsglnAlaIleuAlaPheGluArgTrpAspPhe 149
   ::|||
5075 AAACACACACGACATGACGTATCAACACTAATGTCACTACACGCGTTT 5124
150 ...AspGlnValArgserleuMetGluAsnSerAspArgCys..... 162
   ||| ||||| |||||
5125 ACTGACTGGGTTAACCTACTAGTGAAGAACGCGCTCTGTGTACTACGCA 5174
163 .....GlnAspIleArgAsnleuAlaPheleuGlylleAlaTyraST 177
   ::|||
5175 ATTTAAAAAGAACTTAACAGCAGAACTATTGGCTAAGCTTATGAATC 5224
177 hrlleuLeuArglyleAlaGluIleAlaArglyleArgValys...AspIle 192
   ::|||
5225 TTGATTAATAAATCTACGTGAGCAGCAAGATTCAGGCAAAACTGATTGG 5274
193 SerArg.....ThrAspGlyArgmetleuIleHisIleGlyAr 206
   |||||
5275 TCCAACTGGCCATCTATGACAACTATTGTCAGTGAATTCACG 5324
206 g.....ThrylsTrleuValSerT 213
   |||
5325 TCAATTGAACATCTCCACAAAGGCATTCCACACAGTCAAGTATTGA 5374
213 hrlAla.....GlyValGluLysAla 219
   |||||
5375 CGCAGAAACCTCCGCTAATACCAATGCGAAAGCCGCTGCGATA 5424
220 leuSerleuGlyVal.....ThrylsleuValGluArgTrpI1 232
   |||||
5425 TTAAACCGCGGGGTGACACCAAAAGACTTAACGCGCTTACATGGGTTCT 5474
232 eSerValSerGlyValAlaAspAspProAsnAsnTyrlYleuPheCysArgV 249
   ::|||
5475 GGATGAGTCTCCAGTGTGCTGCTTAAGACACATACTATTATTATCTGATA 5524

```

```

249 alarGLys..... 251
5525 ACCAAAGATCGAAGCAAGAAACATGACATTAATAAACCGTGA 5574
252 .....AsnGlyValAlaAlaProSerAl 259
5575 GATCATATCAATACATATTGATGATGAACCAAGTTCCGCCGACATTA 5624
259 aThSerGlnLeuSer.....Thra 266
5625 AACACACAGATCGCTGACGCTATCGCTACATCCAACTGATATTAAC 5674
266 rGAlaLeu.....GluGly.....IlePheGlnAlaThrHisArg 277
5675 GCGCGTTAAATAATGAGGAGACATACGTGACAAAGTGCACACGCGC 5724
278 LeuIleTyrglyAlaIleAspSerGlyGlnArgTyrlleuAlaTrpSe 294
5725 TCATCTTCACCACTGGGATTAATACAAATTAACGCTACAGCACTGGC 5774
294 rGly.....HisSerA 298
5775 CGGTATGGCTAAACCTTACTATCCAGAAATTACATGATCCGACGC 5824
298 IaArgValGlyAlaAlaArgAspMet.....Ala 307
5825 TACGTATGGCGACAGCAAAATGATGATACGTTGCTGCAATCCATCAGC 5874
308 ArgAlaGlyValSerIleProGluIleMetGlnAla..... 319
5875 CAAGACCAATTAAATATGATACCTAGAACATGCGCTTAATCTTACT 5924
320 .....GlyGlyTrpThrAsnValAsnIleValMetAsnTyrlleA 333
5925 AACATCATTCGAACAGGTGGCTAATCTGAAATCCTCAGCGCCATCAGTC 5974
333 rGAsnLeuAspSerGluTrgIly 340
5975 ACAACATTAAATATGATCAAGCA 5997

seq_name: /cgn2_6/ptodata/2/ina/6a_COMB.seq:US-09-096-942-1
seq_documentation_block:
Sequence 1, Application US/09096942
GENERAL INFORMATION:
APPLICANT: Pollock, Thomas J
APPLICANT: Mikolajczak, Marcia
APPLICANT: Yamazaki, Motohide
APPLICANT: Thorne, Linda
APPLICANT: Armentrout, Richard W
TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
FILE REFERENCE: seq list for appl filed from pro. appl
CURRENT FILING DATE: 1998-06-12
EARLIER APPLICATION NUMBER: US/09/096,942
EARLIER FILING DATE: 1997-06-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 16075
TYPE: DNA
ORGANISM: Xanthomonas campestris
PUBLICATION INFORMATION:
US-09-096-942-1

```

```

alignment_scores:
Quality: 95.50 Length: 317
Ratio: 0.628 Gaps: 18
Percent Similarity: 47.950 Percent Identity: 22.713

```

```

alignment_block:
US-09-662-128a-2 x US-09-096-942-1 ..
Align seq 1/1 to: US-09-096-942-1 from: 1 to: 16075

```

```

54 SerValCysArgSerTrpAlaAlaTrpCysIleuAsnAsnArgIleTr 70
4010 TCGGCTGTGGTTCTGCGGCGCTGCTG.....CGCTG 4044
70 pPheProAlaGluProGluAspValaIArgAspTyrlleuTyrlleuGlnA 87
4045 GTGGCCCGCACGCACTTG...CTGGGTGGCTTCCTCAATCACTCGCGCA 4090
87 IaArgGlyLeuAlaValIys..... 93
4091 CGCAGGGCGGTGATGTCACACGCTGTGTGTAGTGGCGCTGCATCCG 4140
94 .....ThrIleGlnGlnHisLeuGlyGln.....LeuAs 103
4141 GTGATGAAGATCACTCATTAACCTGAGCCGTAAATCCCTGGTGGCATGAA 4190
103 nMetLeu.....HisArgA 108
4191 CATGGTTGGCTATTCCGACCGCGGTACATCGCGCGTGGCGGAAACAG 4240
108 rGserGlyLeuProArgProSerAspSerAsnAlaValSerLeuValMet 124
4241 GCCAGGGCTCCCGGCTGGGTGATCCCATGAGTCACTGATCGATACGTC 4290
125 ArgArgIleArgIysGluAsnVal.....AspAlaGlyGluArg 137
4291 AAGAACCAACAGGTGAGAGGTGTGATGTGATCTCGCTGCCCTGGCGACG 4340
137 gAlaIysGlnAlaLeuAlaPheGluArgTrpAspPheAspGlnValArgS 154
4341 C.....GACCACATCAAGC 4354
154 eLeuMetGluAsnSerAspArg.....CysGlnAspIleArg 166
4355 ACCTGCTGAGCGCTGATGCTACCCGATCAACGTGAACGTGCTGCC 4404
167 AsnLeuAlaPheLeuGlyIleAlaTyraSnThrLeuLeuArgIleAlaG 183
4405 GACCTGTGCACTTGGCCTGTGAACAGCTGCGGACGATCGGACG 4454
183 uIleAlaArgIleArgValIysAspIleSerArgThrAspGlyGlyArgm 200
4455 CGTGGCGGTATCAACCTGGCT.....CAGGGTGGC... 4485
200 eLeuIleHisIleGlyArgTrpThrIleuValSerThrAlaGlyVal 216
4486 .....GTGATCGTGAACACTTCTGCTG.....GTC 4512
217 GluIysAlaLeuSer.....LeuGlyValThrIysIleuValGluArg 230
4513 GCCAAGCGCTGCGAGCAAGATCCTGGCGGTGATGCGCTGATGGCCT 4562
230 gTrp.....IleSerValSerGlyValAlaAspAspProA 242
4563 GTGGCCGCTGATCTGCCCATTCGCTA...GGCGTAAGATGACCTCGC 4609
242 snAsnTyrlleuPheCysArgValaIArgIysAsnGlyValAlaAlaProSe 258
4610 CCGGCGCGGTCTTCCTGCTCAAGCGCGCACGCGCTGGTGGCGGAG 4659
258 rAlaThrSerGlnLeuSerThrArgAlaLeuGluGlyIlePheGlnAlaTr 275
4660 TTCATCATGTTCAAGTTCGCTCGATGC..... 4687
275 hrHisArgLeuIleTyrglyAlaIleAspAspSerGlyGlnArgTyrlleu 291
4688 .....GGGTGATGACGATCATGACACGACGATTC... 4717

```

137 gAlaLysGlnAlaLeuAlaPheGluAyrThrAspPheAspGlnValaLys 154
14341 C..... |||:||||:|
154 erLeuMetGluAsnSerAspArg..... CysGlnAspLeuArg 166
4355 AGCTCTGCTACAGCGCTGGATGGCTACACCGATCAACGTGAAGCTGGTGGCC 4400
167 AsnLeuAlaPheLeuGlyGlyIleAlaTyrrAsnThrLeuLeuArgIleAlaG1 183
4405 GACCTGTTGCACCTTGGCGCTGTGGAACCACTGTGCCGACACATGCGCAG 4445
183 uIleAlaArgIleArvValAspIleSerArgThrAspGlyGlyArgm 200
4455 CGTGGCGGTATCAACTGGCT..... CAGGGTGGC... 4485
200 etLeuIleHisIleGlyAyrThrLysThrLeuValSerThrAlaGlyAl 216
4486GTGGATCGTGACMACACTTGGTG.....GTC 4512
217 GluLysAlaLeuSer.....LeuGlyValThrLysLeuValGluAr 230
4513 GCCAAGCGCGTGCAGACAAAGATCCTGGCGGATTTGGCCTATGGCCT 4562
230 gTTP.....IleSerValSerGlyValAlaAspAspPro 242
4563 GTGGCCCGCTGATCTGGCCATTGCGGTA...GGCCTGAAGATGACCTGCG 4609
242 snAsnTyrlleuPheCysArvValArvLysAsnGlyValAlaAlaProse 258
4610 CCGGCCGGGTCTTCCTCCGTACAGCCGCCACGGCGCTGGGGCGCGAG 4659
258 rAlaThrSerGlnLeuSerThrArgAlaLeuGlnGlyIlePheGluAlar 275
4660 TTCTACATGTTCAAGTTCGCTCGATGC..... 4687
275 hrHisArgLeuIleTyrglyAlaLysAspAspSerGlyGlnArgTyrlleu 291
4688GGGTCCATGACGATCATGGCACACGATTC... 4717
292 AlaTrpSerGlyHisSerAlaArgValAlaAlaAlaArgAspMetAlaAr 308
4718AGCAGCGGCACCAAGACAGACAGACCGGATTAACGCC 4752
308 gaLaglyValSerTlleProGluIleMetGlnAlaGlyGlyTyrrAsn 324
4753 TTGGCAGATTTCC.....TGCGCCGCGACGACCCCTGGACGAGC 4789

Fri May 10 08:35:58 2002

us-09-662-128a-2.rni

Page 14

125 TGGACATGTCAGGATCGCCAGCGCTTTCTGACCATACCTGGAAATG 174
 52 LeuLeuSerValCysArgSerTrpAlaIleArgCysLysLeuAsnAsnAr 68
 |||||||
 175 CTCTCTCCGTTTCCCGGTCGTGGGGGCGATGTCGAAGTTGAATACCG 224
 68 gLySTrPheProIaGluPProGluAspValArgAspTyrLeuLeuTyrL 85
 |||||||
 225 GAATGGTTTCCCGCAGAACCTGAAGATGTTCGCAATTACTTCTATATC 274
 85 euGlnAlaArgGlyLeuAlaValAlaLysThrIleGlnGlnHisLeuGlyGln 101
 |||||||
 275 TTCAAGGCGCGCGCTGCGCATAAACATATCCAGACATTTGGCGCCAG 324
 102 LeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSerAs 118
 |||||||
 325 CTAAACATGCTTCATCGTCGTCGGGCTGCCAGCAAGTGCAGCAG 374
 118 nAlaValSerLeuValMetArgArgIleArgLysGluAsnValAspAlaG 135
 |||||||
 375 TGTGTTCACGTGTATGCGCGCATCCGAAAAAGAAACGTTGATGCCG 424
 135 LysGluArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheAspGln 151
 |||||||
 425 GTGAACGTCGCAAAACAGGCTCTACCGCTTCGACGACTGATTTCCAG 474
 152 ValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAsnLe 168
 |||||||
 475 GTTCCTTACTCATGAGAAATAGCGATCGCTCCGAGATATACGTAATCT 524
 168 uAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGlnIleA 185
 |||||||
 525 GGCATTCTCGGGGATGCTGTATACACCTGTGTACGATAGCCGCAATTTG 574
 185 lAlaArgIleArgValLysAspIleSerArgThrAspGlyArgMetLeu 201
 |||||||
 575 CCAGATCAGGGGTTAAAGATATCTCAGCTACTGACGGGTGGAGAAATGTA 624
 202 lIleHisIleGlyArgThrLysThrLeuValSerThrAlaGlyValGly 218
 |||||||
 625 ATCCATATGGCGCAAGCAAAACGCTGTAGCACCGAGGTGTAGAGAA 674
 218 sAlaLeuSerLeu 222
 |||||||
 675 GGCACCTTACCGCTG 687
 seq_name: gb_est1:AW211323
 seq_documentation_block:
 LOCUS AW211323 431 bp mRNA linear EST 03-DEC-1999
 DEFINITION uc079f01.y1 NCI CGAP Mam3 Mus musculus cDNA clone IMAGE:2648761 5'
 similar to SW:RECR_BP1 P06956 RECOMBINASE CRE. [1] ;, mRNA
 sequence.
 ACCESSION AW211323
 VERSION AW211323.1 GI:6517271
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 431)
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other_ESTS: uc079f01.x1
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.llnl.gov/dbp/image/image.html

MGI:1029213
 Seq primer: -40R from Gibco
 High quality sequence stop: 279.
 Location/Qualifiers

FEATURES

source

1..431
 /organism="Mus musculus"
 /strain="129 - C57/B6 - FVB"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:2648761"
 /clone_id="NCI-CCGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lohar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."
 22, 101 c 111 g 101 t
 BASE COUNT 118 a 111 g 101 t
 ORIGIN

alignment_scores:
 Quality: 652.00 Length: 143
 Ratio: 4.759 Gaps: 0
 Percent Similarity: 95.804 Percent Identity: 90.909

alignment_block:
 us-09-662-128a-2 x AW211323 ..

Align seg 1/1 to: AW211323 from: 1 to: 431

67 AsnArgLysTrpPheProAlaGluProGluAspValArgAspTyrLeuLe 83
 |||||||
 3 AACCGAAATGTTTCCCGCAGAACCTGAAAGATTTGCGATTATCTTCT 52
 83 uTyrLeuGlnAlaArgGlyLeuAlaValLysThrIleGlnGlnHisLeu 100
 |||||||
 53 ATATCTTCAGGCGCGGCTGTGCGAGTAAACATATCCAGCAACATTTGG 102
 100 LysGlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAsp 116
 |||||||
 103 GCCAGCTAAACATGCTCATCTCGCTCGGCTGCCAGCAGCAAGTGAC 152
 117 SerAsnAlaValSerLeuValMetArgArgIleArgLysGluAsnValAs 133
 |||||||
 153 AGCAATGCTGTTCACTGTGTATGCGGCGGATCCGAAAGAAACGTTGA 202
 133 PalagIyGluArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPhe 150
 |||||||
 203 TGCCGGTGAACGTGCNAACAGGCTCTAGCGTTTCAACAGCACTGATTGG 252
 150 spGlnValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArg 166
 |||||||
 253 ACCAGGTGCTTCACTCATGGAATAATAGCAGATCGCTGCAGATATACGT 302
 167 AsnLeuAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaG 183
 |||||||
 303 AATCTGCAATTTCTGGGATTCCTTATACACCGGATCTATATAGCCCG 352
 183 uIleAlaArgIleArgValLysAspIleSerArgThrAspGlyArgM 200
 |||||||
 353 AATGCCAGATCAGGTTAAAGATATCTCAGCGTCTGACGGTGGAGAA 402
 200 eTLeuIleHisIleGlyArgThrLysThr 209
 |||||||
 403 TGTATATCCATATTTGGGAGAGCAAGCAACC 431

seq_name: gb_est2:B1650191

```

seq_documentation_block:
  LOCUS      BI650191          563 bp      mRNA      linear      EST 12-SEP-2001
  DEFINITION  603296283F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5336853 5',
  ACCESSION  BI650191
  VERSION    BI650191.1  GI:15564427
  KEYWORDS   EST.
  SOURCE      house mouse.
  ORGANISM   Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  REFERENCE  1 (bases 1 to 563)
  AUTHORS    NIH-MGC http://mgs.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
  COMMENT    Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-rcmail.nih.gov
              Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM1855 row: f column: 22
              High quality sequence stop: 563.
FEATURES
  source
    1..563
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone_image="5336853"
    /clone_lib="NIH_CGAP_Mam3"
    /tissue_type="tumor, gross tissue"
    /lab_host="DH10B"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
    Average insert 2 kb. Library constructed by Life
    Technologies, catalog #12017-018. Investigators providing
    samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference
    (1999). Note: this is a NCI_CGAP Library."
  BASE COUNT      143 a      133 c      149 g      138 t
  ORIGIN
alignment_scores:
  Quality: 493.00      Length: 97
  Ratio: 5.082      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-09-662-128a-2 x BI650191 ..
  Align seg 1/1 to: BI650191 from: 1 to: 563
254 VALAAlaAProSerAlaThSerGlnLeuSerThrArgAlaLeuGluC1 270
|||||
1 GTTGGCGGCGCATCTGCCACCAAGCATCACTCCGCCCTGGAGG 50
270 YLleheGluAlaThrHisArgLeuIleTyrGlyAlaLysAspAspSerg 287
|||||
51 GATTTTGAAGCAACTCATCATGATTGATTCGCGCTAAGATGACACTGT 100
287 YGlnArgTyrLeuAlaIleTyrSergGlyHisSerAlaArgAlaGlyAla 303
|||||
101 GTCAGAGATACCTGGCTGCTGTGACACAGTCCCGTGGAGCCGCG 150
304 ArgAspMetAlaArgAlaGlyAlaSerIleProGluIleMetGlnAlaG1 320
|||||
151 CGAGATATAGCCCGCGCTGAGATTCAATACCGAGATCATGCAAGCTGG 200
320 YGlyTTrpThrAsnValAsnIleValMetAsnTyrIleArgAsnLeuAsp 337

```

```

seq_name: qb_est1:AW212478
seq_documentation_block:
  LOCUS      AW212478          340 bp      mRNA      linear      EST 03-DEC-1999
  DEFINITION  uc89c09.x1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:2649712 3',
  ACCESSION  AW212478
  VERSION    AW212478.1  GI:6518565
  KEYWORDS   EST.
  SOURCE      house mouse.
  ORGANISM   Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  REFERENCE  1 (bases 1 to 340)
  AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
  JOURNAL    Unpublished (1997)
  COMMENT    Other ESTs: uc89c09.y1
              Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-rcmail.nih.gov
              Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www.dio.llnl.gov/dbp/image/image.html
  MGI:1030164
  Seq primer: -40UP from Gibco
  High quality sequence stop: 158.
FEATURES
  source
    1..340
    /organism="Mus musculus"
    /strain="129 - C57/B6 - FVB/N"
    /db_xref="taxon:10090"
    /clone_image="2649712"
    /clone_lib="NCI_CGAP_Mam3"
    /tissue_type="tumor, gross tissue"
    /dev_stage="10 months"
    /lab_host="DH10B"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
    Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
    Library constructed by Life Technologies, Inc.
    Providing samples: Lohar Hennighausen/Chu-Xia Deng, NIH
    Reference for transgenic model: Xu et al., Nature Genetics
    22, 37-43 (1999)."
  BASE COUNT      75 a      100 c      79 g      86 t
  ORIGIN
alignment_scores:
  Quality: 458.50      Length: 103
  Ratio: 4.776      Gaps: 1
  Percent Similarity: 93.204      Percent Identity: 89.320
alignment_block:
  US-09-662-128a-2 x AW212478/rev ..
  Align seg 1/1 to reverse of: AW212478 from: 1 to: 340
248 ArgValaArgLysAsnGlyValaAlaAProSerAlaThSerGlnLeuS 264
|||||
340 CGGGTCAGAAAAAATGTGTTCGCCCATCTGCCACAGGCGCAAGCTGCA 291

```

```

264 rThrArgAlaLeuGluglyIlePheGluAlaThrHisArgLeuIleTyrG 281
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
290 CTCGGG...CTGAGAGGATTTTGAAGCACTCATCGATTGATTACG 244
281 lAlaAlaYsAspSperGlyGlnArgTyrLeuAlaTrpSerGlyHisSer 297
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 GCGTAAGCATGATCTGTGTAGAGATACCTGGCTGTGTGACACACAGT 194
298 AlaArgValGlyAlaAlaArgAspMetAlaArgAlaGlyValSerIlePr 314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 GCCCTGTGTGAGCGCGGAGATATGCCCCGCTGTGATTCATACACC 144
314 GGUlMetGlnAlaGlyIleTyrThrAsnValAsnIleValMetAsnT 331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
143 GGAGATCATGCAAGCTGTGCTGACCAATGTAATATTGTCTCATGAACT 94
331 YrIleArgAsnLeuAspSerGluThrGlyAlaMetValArgLeuLeuGlu 347
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
93 ATATCCGTAAACCTGGATGTAACAGGGGCAATGTGCTGTGTGGA 44
348 AspGlyAsp 350
      ||||| |||||
43 GATGGCGAT 35
seq_name: gb_est1:AW211972
seq_documentation_block: 338 bp mRNA linear EST 03-DEC-1999
LOCUS AW211972
DEFINITION u09f01.x1 NCI CGAP Mam3 Mus musculus CDNA clone IMAGE:2648761 3'
      similar to SW:RCCR_BPPI P06956 RECOMBINASE CRE. [1] ;, mRNA
      sequence.
ACCESSION AW211972
VERSION AW211972.1 GI:6517961
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 338)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
      Tumor Gene Index
      Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Other ESTs: u079f01.y1
      Email: cgapbs-remail.nih.gov
      Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
      CDNA Library Preparation: Life Technologies, Inc.
      DNA Sequencing by: Washington University Genome Sequencing Center
      Clone distribution: NCI-CGAP clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      www-bio.llnl.gov/dbtrp/image/image.html
FEATURES
      source
      MGI:1029213
      Seq primer: -40UP from Gibco
      High quality sequence stop: 151.
      Location/Qualifiers
          1..338
              /organism="Mus musculus"
              /strain="129 - C57/B6 - FVB/N"
              /db_xref="taxon:10090"
              /clone_image="IMAGE:2648761"
              /clone_lib="NCI CGAP Mam3"
              /tissue_type="tumor, gross tissue"
              /dev_stage="10 months"
              /lab_host="DH10B"
              /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
              Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
              Library constructed by Life Technologies. Investigators
              providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
              Reference for transgenic model: Xu et al., Nature Genetics

```

```

BASE COUNT      22, 37-43 (1999)..
ORIGIN          78 a 93 c 82 g 85 t

```

```

alignment_scores:
  Quality: 397.50      Length: 103
  Ratio: 4.321        Gaps: 2
  Percent Similarity: 89.320      Percent Identity: 87.379

```

```

alignment_block:
US-09-662-128A-2 x AW211972/rev ..

```

```

Align seg 1/1 to reverse of: AW211972 from: 1 to: 338

```

```

249 ValArgLysAsnGlyValAlaAlaPro.SerAlaThrSerGlnLeuSert 265
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
338 GTCCACAAATAATGGTGTGTGCGCCATAGCTACCTG...CTGCATCA 292
265 hrArgAlaLeuGluglyIlePheGluAlaThrHisArgLeuIleTyrGly 281
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 CTCGGGCGCTGGAAAGGAT.TTCTAAGCCACTCATCGATTGATTACGCG 243
282 AlaAlaYsAspSperGlyGlnArgTyrLeuAlaTrpSerGlyHisSerAl 298
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
242 GCTAAGGATGACTTGTGTAGAGATACCTGCTGTGTGACACACAGTC 193
298 aArgValGlyAlaAlaArgAspMetAlaArgAlaGlyValSerIleProG 315
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
192 CCGTGTGAGAGCGCGGTTCATATGCCCCGCTGTGATTCATACCGG 143
315 lAlMetGlnAlaGlyIleTyrThrAsnValAsnIleValMetAsnTyr 331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
142 AGATCATGCAAGCTGTGCTGACCAATGTAATATTGTCTATACACTAT 93
332 lLeArgAsnLeuAspSerGluThrGlyAlaMetValArgLeuLeuGluAs 348
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
92 ATCCGTAAACCTGGATGTAACAGGGGCAATGTGCTGTGTGGA 43
348 pGlyAsp 350
      ||||| |||||
42 TGCGCGAT 36
seq_name: gb_est1:AW210918
seq_documentation_block: 343 bp mRNA linear EST 03-DEC-1999
LOCUS AW210918
DEFINITION u089c09.y1 NCI CGAP Mam3 Mus musculus CDNA clone IMAGE:2649712 5'
      similar to SW:RCCR_BPPI P06956 RECOMBINASE CRE. [1] ;, mRNA
      sequence.
ACCESSION AW210918
VERSION AW210918.1 GI:6516858
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 343)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
      Tumor Gene Index
      Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Other ESTs: u089c09.x1
      Email: cgapbs-remail.nih.gov
      Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
      CDNA Library Preparation: Life Technologies, Inc.
      DNA Sequencing by: Washington University Genome Sequencing Center
      Clone distribution: NCI-CGAP clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      www-bio.llnl.gov/dbtrp/image/image.html

```

MGI:1030164
Seq primer: -40RP from Gibco
High quality sequence stop: 222..
Location/Qualifiers
1..343

FEATURES
source
/organism="Mus musculus"
/strain="129 - C57/B6 - FVBN"
/db_xref="taxon:10090"
/clone_1lb="IMAGE:2649712"
/clone_1lb="MCI CGAP Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."
BASE COUNT 95 a 77 c 85 g 84 t 2 others
ORIGIN

alignment_scores:
Quality: 393.50 Length: 115
Ratio: 3.896 Gaps: 2
Percent Similarity: 87.826 Percent Identity: 80.000

alignment_block:
US-09-662-128a-2 x AM210918 ..

Align seg 1/1 to: AM210918 from: 1 to: 343

```

107 ArgArgSerGlyLeuProArgProSerAspSerAsnAlaValSerLeuVa 123
2 CGTCGGTCGGGGCTGCCACGACGACAGACGACATCTCTTCTACCTGCT 51
123 lMeArArgIleArgIlysgIusnValAspAlaIlyGluArgAlaIlysg 140
52 TATCCGGCGGATCCGAAAGAAACGTTGATGCCGGTGAAACGTGCAAAAC 101
140 lAlaIleAlaIlePheGluArgThrAspPheAspGlnValArgSerLeuMet 156
102 AGGCTCAGCGCTTCGAACGACATCTTCCAC.CAGGTCGTTCACTCATG 150
157 GluAsnSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyI 173
151 GAAATATAGCATCGCTGCCAGATATACGTAATCTGCGCATTTGGGGGAT 200
173 eAlaArgAsnThrLeuLeuArgIleAlaIleAlaIleAlaArg... 188
201 TTGCTTATTAACCCCTGTACGTATACCCGAAATTCGACAGATCCANGG 250
189 ValIysAspIleSerArgThrAspGlyIArgMetLeuIleHisIleG 205
251 TTAAGATCTCTTACGCTTGACGTGGAGAAAAAATGTAATCCATATG 300
205 YArgThrIlysthrLeuValSerThrAlaGlyValGluIlysaIa 219
301 CAGAAACAAAACCCCTGTATACCCGCCACGCTGTAGAAAGGCC 343

```

seq_name: gb_gss:FR0013439

seq_documentation_block:
LOCUS FR0013439 288 bp DNA linear GSS 18-SEP-1997
DEFINITION F.rubripes GSS sequence, clone 122014bg11, genomic survey sequence.
ACCESSION AL004689
VERSION AL004689.1 GI:2450259
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes.
ORGANISM Takifugu rubripes.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

REFERENCE
1 (bases 1 to 288)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umranta,Y., Williams,G. and Brenner,S.
TITLE Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10 1SB. Email: biolhelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V-type: phagemid
PRIMER: KS
DESCR: One pass dye-terminator sequencing of cosmid cloned genomic sequence.

FEATURES
source
Location/Qualifiers
1..288
/organism="Takifugu rubripes"
/db_xref="taxon:31033"
/clone_1lb="cosmid 122014"
/clone_1lb="122014bg11"
BASE COUNT 72 a 70 c 70 g 73 t 3 others
ORIGIN

alignment_scores:
Quality: 261.50 Length: 78
Ratio: 4.218 Gaps: 6
Percent Similarity: 79.487 Percent Identity: 78.205

alignment_block:
US-09-662-128a-2 x FR0013439 ..

Align seg 1/1 to: FR0013439 from: 1 to: 288

```

1 MetProIlyslsIlysrGlyValSerAsnLeuLeuThrValHisGlnAs 17
54 ATGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 103
17 neuProAlaIleuProValAspAlaThrSerAspGluValAlaGlyAsnL 34
104 TTTCCTGCTATACCGTGCATGATGCAACAGATGATGATGATGATGATGAT 153
34 euMetAspMet.PheArgAspArgGln...AlaPheSerGlnHisThrTr 49
154 TGATGACATGTTTCAGAGGATTCGCCAGGCGTTTTCCTGACACATACCTG 203
49 pIys...MetLeuLeuSerValCysArg...SerTrpAlaAlaTrp.Cys 63
204 GAAATATGCTTCCTGTTCCGTTTCCCGGTCGCTGGCGGCGCATGTTGC 253
64 LysLeuAsnAsnArgIlystrPhePro 72
254 AANTTTGAATTAACCCGAAATGCTTCCCC 281

```

seq_name: gb_est2:B1556329

seq_documentation_block:
LOCUS B1556329 836 bp mRNA linear EST 05-SEP-2001
DEFINITION 60323744F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5290619 5', mRNA sequence.
ACCESSION B1556329
VERSION B1556329.1 GI:15443643
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 836)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

seq_name: gb_gss:AZ204748

seq documentation block:

LOCUS AZ204748 869 bp DNA linear GSS 31-AUG-2000
DEFINITION SP_0097_B2_A06_T7A Strongylocentrotus purpuratus, purple sea urchin,
, sperm genomic BAC library Strongylocentrotus purpuratus genomic
clone Plate=97 Col=12 Row=B, DNA sequence.
ACCESSION AZ204748
VERSION AZ204748
KEYWORDS GSS.

SOURCE
ORGANISM Strongylocentrotus purpuratus.
Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoidea;
Strongylocentrotidae; Strongylocentrotus.
REFERENCE Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
AUTHORS Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray
,G.A., Ettensohn,C.A., Demach,H., Britten,R.J., Davidson,E.H. and
Hood,L.
TITLE A sea urchin genome project: Sequence scan, virtual map, and
additional resources
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)
MEDLINE 20402566

COMMENT

Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu

Plate: 97 row: B column: 12
Seq primer: T7
Class: BAC ends

FEATURES

High quality sequence stop: 869.

Source

Location/Qualifiers

1. 869
/organism="Strongylocentrotus purpuratus"
/db_xref="taxon:7668"
/clone_lib="Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BACs 6; BAC Clones in E-Coli DH10B"

BASE COUNT 285 a 149 c 199 g 236 t

ORIGIN

alignment_scores:

Quality: 99.50 Length: 234
Ratio: 0.858 Gaps: 11
Percent Similarity: 49.573 Percent Identity: 24.786

alignment_block:

US-09-662-128a-2 x AZ204748 ..

Align seg 1/1 to: AZ204748 from: 1 to: 869

```

82 LeuLeuTYrLeuGlnAlaArgGlyLeuAlaValysThrIleGlnIlnH 98
   :::::|||||
75 GGGCTGTATACGCCCATTAAGTGGGCTACTACTCTA..... 110
   :::::|||||
98 sLeuGlyIlnLeuAsnMetLeuHisArgArgSerGlyLeuProArgPro 115
   :::::|||||
111 .....GAGGGTTCCAGGCGATCCACAG.... 131
   :::::|||||
115 erAspSerAsnAlaValSerLeuValMetArgAlaGlyLeuGlnAsn 131
   :::::|||||
132 .....ACAAATGTTTGTAAACAATT...TGTTCGAAAT 161
   :::::|||||
132 ValAspAlaGlyGlnArgAlaValGlnAlaLeuAlaPheGlnArgThrAs 148
   :::::|||||
162 CTGCCAAACGGCAATGCACGCAAAAAGTAAACAGAAAGATCCAGTTAC 211
   :::::|||||
148 pPheAspGlnValArgSerLeuMetGluAsnSerAspArgCysGlnAsp 165
   :::::|||||
212 TTCAGATCAAGTTAAAGACGATCTAAGAACATGTTCAATACAGAGAC. 260
   :::::|||||
165 leuArgAsnLeuAlaPheLeuGlyIlnAlaTyrAsnThrLeuLeuArgIle 181
   :::::|||||
261 ..GCAATTTTGTCTTTCACGGG.....TTTCTTAGATAT 293
   :::::|||||
182 AlasGluIlnLeuAlaArgIleArgValysAspIleSerArgThrAspGly 198
   :::::|||||
294 GAGGAACACAGATCTTGAAGGTGTAATGATGTGCAATTTATGAT...GA 340
   :::::|||||
198 YArgMetLeuIleHisIleGlyArgThrIlyThrLeuValSerThrAlaG 215
   :::::|||||
341 TTAATATTCTTAATATTCCTAAGAGTAAGACAGATCAGTAAAGCAAG 390
   :::::|||||
215 lYValGlnLysAlaLeuSerLeuGlyValThr.....Lys 226
   :::::|||||
391 GTAAATAGTGAATCATTAAGTCCGGGTGTGACAGAACTTGTCGGTCAAA 440
   :::::|||||
227 LeuValGlnArgTrpIleSerValSerGlyValAlaAspAspProAsnAs 243
   :::::|||||
441 ATGCTAAAGAAATATCAATGATGCGCTCGGATTTCTTTAGATTTCCAAATCA 490

```

seq_name: gb-gss:BH093306

seq_documentation_block:

LOCUS BH093306 787 bp DNA linear GSS 18-JUL-2001
DEFINITION RPCI-24-270M12.TV RPCI-24 Mus musculus genomic clone RPCI-24-270M12
, DNA sequence.
ACCESSION BH093306
VERSION BH093306.1 GI:14913211
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 787)
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Atinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shwartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other-GSSs: RPCI-24-270M12.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
plate: 270 row: M column: 12
Seq primer: T7
Class: BAC ends.

TITLE
JOURNAL
COMMENTFEATURES
Source

Location/Qualifiers

1. 787
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="RPCI-24-270M12"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pPARBAC1. Site 1: BamHI. Site 2: BamHI;
RPCI-24 Mouse BAC library produced by Pieter de Jong. The
library was cloned in the pPARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 254 a 161 c 178 g 194 t

ORIGIN

alignment_scores:

Quality: 96.00 Length: 97
 Ratio: 1.778 Gaps: 3
 Percent Similarity: 55.670 Percent Identity: 28.866

alignment_block:

US-09-662-128a-2 x BH093306 ..

Align seg 1/1 to: BH093306 from: 1 to: 787

```

3 LysIlyslsArglyValSerAsnLeuThrValHisGln.AsnLeuP 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
417 AAAAAAGTGGCTAAATTGCTGGCGCAGTGGTGGTGCACACTTTAAATCC 466
19 r0AlaLeuProValAlaSpAlaThrSerAsp.....GluValArgLys 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
467 CCGCACTCAAGAGGTAGAGGAGGAGCATCTTGAGTTCAAGGCCACTT 516
33 AsnLeuMetAspMetLpheArgAspArgGlnAlaPheSerGluHisThr.. 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 GGTCTACAAAGCAGAGTTCAGAGCAGCCAGTGTACACAGAGAAACCTT 566
49 .....TrpLysMetLeuLeuSerValLysArgSerT 59
567 AATGATATACAAATGAGGCTGAGCTAGCCCTGCAATATGATGACAGAT 616
59 rPAlaAlaIrrpCysLysLeuAsnAsnArgLysTrpPheProAlaGluPro 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
617 GGGCAGACTGGTGGTTCATGACAGTCCCAAAACACACAGCAGCATCTT 666
76 GluAspValAlaArgAspTyrLeuLeuTyrLeuGlnAlaArg 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
667 GCACCCCAATGAGCACCTGCTGTGGCCCGCTGGGAGAG 705

```

seq_name: qb_gss:AG083458

seq_documentation_block:

LOCUS AG083458 1287 bp DNA linear GSS 03-NOV-2001
 DEFINITION Pan troglodytes DNA, clone: PTB-080013.R, genomic survey sequence.
 ACCESSION AG083458
 VERSION AG083458.1 GI:16635260
 KEYWORDS GSS; GSS (genome survey sequence).
 SOURCE Pan troglodytes male lymphoblast DNA, clone:lib:PTB Chimpanzee Male
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
 AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
 Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE
 JOURNAL BAC end sequences of Library PTB
 REFERENCE
 AUTHORS Unpublished
 2 (bases 1 to 1287)
 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
 Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE
 JOURNAL Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Tensho-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimpesgsc.riken.go.jp, Fax:81-45-503-9170)
 Tel:81-45-503-9111, Fax:81-45-503-9170
 Clones are derived from the chimpanzee BAC library PTB. This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS

Sequencing: M13rev

LIBRARY

Vector : PK5145
 R.site 1 : SacI
 R.site 2 : SacI
 Location/Qualifiers
 1..1287

/organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="PTB-080013.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"
 BASE COUNT 222 a 569 c 379 g 99 t 18 others
 ORIGIN

alignment_scores:

Quality: 96.00 Length: 226
 Ratio: 0.865 Gaps: 10
 Percent Similarity: 49.115 Percent Identity: 25.664

alignment_block:

US-09-662-128a-2 x AG083458/rev ..

Align seg 1/1 to reverse of: AG083458 from: 1 to: 1287

```

142 LeuAlaPheGluArgThrAspPheAspGlnValArgSerLeuMetGluAs 158
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
873 CTGTGCTGCTGCGCGGTGACATCGGGCGCGCGCGCG.....TCGCG 830
158 nSerAspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyIleAlaT 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
829 GCGGAGCGCGGTGGCGACGTGCGGCTTCGCGGAGCTGGGCGCTCTG. 781
175 rAsnThrLeuLeuArgIleAlaGlnIleAlaArgIleArgVal..... 189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
780 .....GCGCTTCGCGCGGTGCGCGCGGTGCGGTGCGGTGCGGTGTC 736
190 .....LysAspIleSerArgThrAspG 197
735 GTGGCAGTGTGCGGTGCGCGGAGTACGCGGCGCGGTGCGCGGCGGCG 686
197 yGluArgMetLeuIleHisIleGlyArgThrLysThrLeuValSerThra 214
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
685 CAGCAGG.....GTGGCGCGCGCGCGCATGTGTCGTGCGCGGAGG 642
214 laGlyValGluLysAlaLeuSerLeuGlyValThrLysLeuValGluArg 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
641 CCGGGGTGGCGCGGTGCTGCTACTC..... 616
231 TrpIleSerValSerGlyValAlaAlaAspProAsnAsnTyrLeuPheCy 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
615 ...ATCAGTCTGCGCGCGCGCTGTGCGACTATCGGTAGTGGTCTGCG 569
247 sarValAlaArgLysAsn.....GlyValA 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
568 GCGCGCGCGCGCGCGCGCTGCGGGGCGGAGGCGCGCTGCGGTCCGG 519
255 laAlaProSerAlaThrSerGlnLeuSerThrArgAlaLeuGluIlyle 271
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
518 GCGGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCACAGCGCGCGCAT 469
272 PheGluAlaThrHisArgLeuIleTyrGlyAlaLysAspAspSer..... 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
468 GCGGCGCGGGGGGTGCGGTAAATGTGCGGCGTGTGAGGCGCGGGG 419
287 .GlyGlnArgTyrLeuAlaIrrpSerGly..... 295
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 GGGGTGCGGTGGTGGCGCGGTGCGGCGAGTGGCGGCGGGCGGGTGG 369
296 .....HisSerAlaArgValGlyAla 302
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
368 CACCTACAGTCTGCATTCGCGGGCGCTCACATACACTGACAGTGGGGCC 319
303 AlaArgAspMetAlaArgAlaGlyValSerIlePro.....GluI 316
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
318 GCGCGG.....GTGGCAGCGCGGTGCGGCGGTGCTGTGGCGTGGCGGT 272
316 emetGlnAlaGlyIlyTrpThrAsnVal 325

```

```

      :::: ||||| ::::|
271 GGTGCGGCGGGGGCTGGTGTGTCG 244
seq_name: gb_est2:BF797551

seq_documentation_block:
LOCUS      BF797551              778 bp    mRNA    linear    EST 12-JAN-2001
DEFINITION  602257335F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4340720 5',
            mRNA sequence.
ACCESSION   BF797551
VERSION     BF797551.1 GI:12102605
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 778)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM953 row: e column: 09
            High quality sequence stop: 749.
            Location/Qualifiers
              1..778
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4340720"
                /clone_lib="NIH_MGC_85"
                /tissue_type="lymphoma, cell line"
                /note="Organ: lymph. Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 1.867 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."

BASE COUNT      159 a      225 c      230 g      164 t
ORIGIN
alignment_scores:
  Quality:      94.50      Length:      198
  Ratio:        0.936      Gaps:          8
  Percent Similarity: 51.010      Percent Identity: 27.273

alignment_block:
US-09-662-128a-2 x BF797551/rev ..
Align seg 1/1 to reverse of: BF797551 from: 1 to: 778

160 AsparGcGlnAspIleArg.....AsnLeuAl 169
|||||
672 GACAGGTGCCAGGACACGCTGGGCTCGCTCTACAGACAACTCTAA 623
169 apheLeuGlyTleatyrAsnThrLeuLeuArgIlealGluIleAla 186
||| ::::|
622 AGGAGCTGCCAAGCTCGAGACGGAGATGACGTGCGAAGCTCAGACA 573
186 rgIleArgValIysAspIleSerArgThrAspGlyGlyArgMetLeu... 201
|||||
572 GGTGATGAGTTCATCCCTGTCAATGTCAAGCTCCCGAGATGTTGTCG 523
202 .....IleHisIleGlyArgThrIleuValSerThr..... 213
|||||
522 ATGAGCTGTTCAATCTCAGACGCACTAAGCCGTGTGTCTCGCCCTTCC 473

```

```

214 .....AlaGlyValGlyValAlaLeuS 221
472 CGTAGGCACTTACACCGCGCTCAGGTTCTTCTTCAAGGTTCCCT 423
221 erLeuGlyValThrLysLeuValGluArg...TrpIleSerValSerGly 236
|| ::::|
422 CATCATCAAAAGTCAAGATGCGGAGGACATTAATGGACTTGAATGTCGGC 373
237 ValAlaAspAspProAsnAsnTyrLysPheCysArgValArgLysAsnG1 253
|||||
372 GTGGCTGTGTCACTGAACACA.....CTGAGAGATCCAGAGAACTCTTC 329
253 Y.....ValAlaAlaProSerAlaThrSerGlnLeuSerThrA 266
||| ::::|
328 AAGGCTAAGGCTGTCTTGGCTGGGATGTGGAGAAAGCCCTGCAGATTC 279
266 rglAlaLeuGluGlyIlePheGluAlaThrHisArgLeuIleTyrGlyAla 282
|||||
278 GCTCTTGAAGGGGTTGGCTTGAGCTCTGGAAGGCTGAGAATCTGCTCG 229
283 Lys.....AspAspSerGlyAlaArgTyrLeuAlaTr 293
|||
228 AAGGCACTTGTGCCGAGTAGCAGATCCACGCTCCGCTCTCTGCGG 179
293 pSer.GlyHisSerAlaArgValGlyAlaAlaArgAspMetAlaArgAla 309
||| ::::|
178 AAGCAGCTCAACAAACCGCTGTGGGTAGGAGATCTCTGCTTCGTCA 129
310 GlyValSerIleProGluIleMetGlnAlaGlyGlyTyrThr 323
|||||
128 GGAGCTCAAGTCTGTGATCTGCGCCAGCAGCTCTCTTGACA 87

seq_name: gb_est2:BG829464

seq_documentation_block:
LOCUS      BG829464              894 bp    mRNA    linear    EST 22-MAY-2001
DEFINITION  602763759F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899062 5',
            mRNA sequence.
ACCESSION   BG829464
VERSION     BG829464.1 GI:1417051
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 894)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM1790 row: e column: 15
            High quality sequence stop: 866.
            Location/Qualifiers
              1..894
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4899062"
                /clone_lib="NIH_MGC_42"
                /tissue_type="epithelioid carcinoma cell line"
                /note="Organ: pancreas; Vector: pORF7; Site_1: XhoI;
                Site_2: EcoRI; cDNA made by oligo-dT priming.
                directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Size-selected >500bp

```

for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library.

BASE COUNT 194 a 253 c 255 g 191 t 1 others

alignment_scores: 94.00 Length: 286
Quality: 0.676 Gaps: 14
Ratio: 0.676
Percent Similarity: 48.601 Percent Identity: 24.476

alignment_block:
US-09-662-128A-2 x BG829464/rev ..

Align seg 1/1 to reverse of: BG829464 from: 1 to: 894

```

52 LeuLeuSerValCysArgSerTrpAlaAlaTrpCysLysLeuAsnAsnAr 68
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
759 CTGCACACGCGCAGCTCCAGCTGGCGCCGCGCCAGCTGGCAAC.. 712
68 GlyTrpPheProAlaGluProGluAspValArgAspTyrLeuLeuTyrL 85
712 ..... 712
85 euGlnAlaArgGlyLeuAlaValLysThrIleGlnIleAsnGlyGln 101
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
711 .....ACAGCTTGACATTGGCGCAGCTAGCAAGAGTTCCT 670
102 LeuAsnMetLeuHisArgArgSerGlyLeuProArgProSerAspSerAs 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 TGG.....ACAGGTTGCCA.....GACAC 650
118 nAlaValSerLeuValMet.ArgArgIleArgLysGluAsnValAspAla 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
649 CACGCGGGGCTGCTGCACAGACACTTAAGAGACCTGCGCAAGTCT 600
135 GlyLysArgAlaLysGlnAlaLeuAlaPheGluArgThrAspPheAspG 151
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
599 GGAGAACGG.....GAGATGACGCTGCTGCACTC 571
151 nValArgSerLeuMetGluAsnSerAspArgCysGlnAspIleArgAsnL 168
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
570 AGAGAGCTGATGATGCTCCATCCCTGCAATGCACAGCTCCACAGAGT 521
168 euAlaPheLeuGlyIleAlaTyrAsnThrLeuLeuArgIleAlaGluL 184
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
520 TGTGATGAGCTGCTTCATCTCAGACGCACTAAGCCGTGTCTCGCC 471
185 AlaArgIleArgValLysAspIleSerArgThrAspGlyArgMetLe 201
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
470 TCTCCCTGAGG...CAGTTCCACACCGCGCTC..... 441
201 uIleHisIleGlyArgThrLysThrLeuValSerThrAlaGlyValGluL 218
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
440 .....AGTCTCTCTGTTCA 425
218 yAlaAlaLeuSerLeuGlyValThrLysLeuValGluArg...TrpIleSer 233
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
424 AGGTTCCGTCATCTCAAGCTCAAGATGGGAAAGCATTAATGAGACTGG 375
234 ValSerGlyValAlaAspAspProAsnAsnTyrLeuPheCysArgValAr 250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
374 ATGCTGCGCGTGTGTCTACATGACACACA.....CTGAGAGATCCAG 331
250 gLysAsnGly.....ValAlaAlaProSerAlaThrSerGln 263
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
330 GAAGTCCCTCAAGCTAAGCTGTCTTGGTGGGATGTGGAAGACCC 281
263 euSerThrArgAlaLeuGluGlyIlePheGluAlaThrHisArgLeuLe 279
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

280 TGCAGATGCTCTCTGAGAGGGTTGGCCTTGAGCTGTGAAGGCTGAGA 231
280 TyrGlyAlaLys.....AspAspSerGlyGlnArgTyr 290
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 ATCTGCTCGAAGGCGACTTGTGCCGCAAGTGCAGCTCAGCTCCGC.. 183
290 rLeuAlaTrpSerGlyHisSerAla...ArgValGlyAlaAlaArg..Asp 305
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 .TGCTCTGGGGAAGCAGCTCACAANAACCGCTGTGGCTAGAGAGATCT 134
306 MetAlaArgAlaGlyValSerIleProGluIleMetGlnAlaGlyTyr 322
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 CCTGCTTCGTCAAGAAAGTCAAGTCCCTGTACTGCGCCAGCAAGCTCTTG 84
322 pThr 323
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 GACA 80

```

seq_name: gb_est2:BE889912

seq documentation block: 688 bp mRNA linear EST 20-OCT-2000
LOCUS BE889912
DEFINITION 601512164F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3913598 5',
mRNA sequence.
ACCESSION BE889912
VERSION BE889912.1 GI:10347709
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 688)
NIH-MGC http://mhc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA sequencing by: Incyte Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRAM9733 row: p column: 15

High quality sequence stop: 688.
Location/Qualifiers

FEATURES
source
1..688
location/Qualifiers

1..688
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3913598"
/clone_id="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 145 a 205 c 190 g 148 t

alignment_scores: 93.50 Length: 204
Quality: 0.908 Gaps: 10
Percent Similarity: 50.490 Percent Identity: 26.961

alignment_block:
US-09-662-128A-2 x BE889912/rev ..

Align seg 1/1 to reverse of: BE889912 from: 1 to: 688

```

160 AspArgCysGlnAspIleArgAsnLeuAlaPheLeuGlyIleAlaTyrAs 176
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Fri May 10 08:35:59 2002

us-09-662-128a-2.rst

Page 12

407 TCTGTACATGCGTTATATCCGCTAAATATCAGGTGCACT 448
